



Browning, T. D., & Matthiesen, L. (2017). Norm forms for arbitrary number fields as products of linear polynomials. *Annales Scientifiques de l'École Normale Supérieure*, 50(6), 1383-1446.  
<https://doi.org/10.24033/asens.2348>

Peer reviewed version

Link to published version (if available):  
[10.24033/asens.2348](https://doi.org/10.24033/asens.2348)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Societe Mathematique de France at [http://smf4.emath.fr/Publications/AnnalesENS/4\\_50/html/ens\\_ann-sc\\_50\\_1383-1446.php](http://smf4.emath.fr/Publications/AnnalesENS/4_50/html/ens_ann-sc_50_1383-1446.php) . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# NORM FORMS FOR ARBITRARY NUMBER FIELDS AS PRODUCTS OF LINEAR POLYNOMIALS

T.D. BROWNING AND L. MATTHIESEN

ABSTRACT. Given a number field  $K/\mathbb{Q}$  and a polynomial  $P \in \mathbb{Q}[t]$ , all of whose roots are in  $\mathbb{Q}$ , let  $X$  be the variety defined by the equation  $\mathbf{N}_K(\mathbf{x}) = P(t)$ . Combining additive combinatorics with descent we show that the Brauer–Manin obstruction is the only obstruction to the Hasse principle and weak approximation on any smooth and projective model of  $X$ .

RÉSUMÉ. Étant donné un corps de nombres  $K/\mathbb{Q}$  et un polynôme  $P \in \mathbb{Q}[t]$ , dont toutes les racines sont dans  $\mathbb{Q}$ , soit  $X$  la variété définie par l'équation  $\mathbf{N}_K(\mathbf{x}) = P(t)$ . En combinant la combinatoire additive avec la descente, nous montrons que l'obstruction Brauer–Manin est le seul obstacle au principe de Hasse et à l'approximation faible sur un modèle projectif et lisse de  $X$ .

## CONTENTS

1. Introduction	2
2. Algebraic number theory	5
3. Technical tools	11
4. Norm forms modulo $p^m$	14
5. Counting points on systems of norm form equations	18
6. $W$ -trick and non-correlation with nilsequences	27
7. Majorants for positive multiplicative functions	32
8. Construction of the majorant	40
9. The majorant is pseudorandom	52
10. Conclusion of the proof	62
References	65

---

*Date:* September 7, 2016.

*Key words and phrases.* additive combinatorics, Brauer–Manin obstruction, descent, Hasse principle, norm forms, weak approximation.

2010 *Mathematics Subject Classification.* 14G05 (11B30, 11D57, 11N37, 14D10).

## 1. INTRODUCTION

Let  $K/\mathbb{Q}$  be a finite extension of number fields of degree  $n \geq 2$  and fix a basis  $\{\omega_1, \dots, \omega_n\}$  for  $K$  as a vector space over  $\mathbb{Q}$ . We will denote by

$$\mathbf{N}_K(x_1, \dots, x_n) = N_{K/\mathbb{Q}}(x_1\omega_1 + \dots + x_n\omega_n)$$

the corresponding norm form, where  $N_{K/\mathbb{Q}}$  denotes the field norm. The objective of this paper is to study the Hasse principle and weak approximation for the class of varieties  $X \subset \mathbb{A}^{n+1}$  satisfying the Diophantine equation

$$P(t) = \mathbf{N}_K(x_1, \dots, x_n), \quad (1.1)$$

where  $P(t)$  is a product of linear polynomials all defined over  $\mathbb{Q}$ . If  $r$  denotes the number of distinct roots of  $P$ , then  $P$  takes the form

$$P(t) = c^{-1} \prod_{i=1}^r (t - e_i)^{m_i}, \quad (1.2)$$

for  $c \in \mathbb{Q}^*$ ,  $m_1, \dots, m_r \in \mathbb{Z}_{>0}$  and pairwise distinct  $e_1, \dots, e_r \in \mathbb{Q}$ .

We let  $X^c$  be a smooth and projective model of  $X$ . Such a model  $X^c$  need not satisfy the Hasse principle and weak approximation, as has been observed by Coray (see [5, Eq. (8.2)]). Specifically, when  $P(t) = t(t-1)$  and  $K$  is the cubic extension  $\mathbb{Q}(\theta)$ , obtained by adjoining a root  $\theta$  of  $x^3 - 7x^2 + 14x - 7 = 0$ , then the set  $X^c(\mathbb{Q})$  is not dense in  $X^c(\mathbb{Q}_7)$ . It has, however, been conjectured by Colliot-Thélène (see [4]) that all counter-examples to the Hasse principle and weak approximation for  $X^c$  are accounted for by the Brauer–Manin obstruction.

This conjecture covers the more general case where  $X$  arises from an equation of the form (1.1), but the ground field may be an arbitrary number field  $k$  instead of  $\mathbb{Q}$  and the polynomial need not factorise completely over  $k$ . In this more general setting the problem of establishing Colliot-Thélène’s conjecture has been addressed under various assumptions on the extension  $K/k$  and upon the polynomial  $P(t)$ . Thus the conjecture is now known to be true for Châtelet surfaces ( $[K : k] = 2$  and  $\deg(P(t)) \leq 4$ ) by work of Colliot-Thélène, Sansuc and Swinnerton-Dyer [10, 11], a family of singular cubic hypersurfaces ( $[K : k] = 3$  and  $\deg(P(t)) \leq 3$ ) by work of Colliot-Thélène and Salberger [5], the case where  $K/k$  is arbitrary and  $P(t)$  is split over  $k$  with at most two distinct roots (see [9, 23, 31, 34]) and the case where  $K/\mathbb{Q}$  is arbitrary and  $P(t)$  is an irreducible quadratic polynomial over  $\mathbb{Q}$  (see [1, 13]). Finally, if one assumes Schinzel’s hypothesis, then it is true for  $K/k$  cyclic and  $P(t)$  arbitrary, by work of Colliot-Thélène, Skorobogatov and Swinnerton-Dyer [12].

Suppose now that  $k = \mathbb{Q}$  and  $P$  is given by (1.2). Until recently, Colliot-Thélène’s conjecture was only known to hold unconditionally when  $r \leq 2$ . When  $r \leq 1$ , the variety  $X$  is a principal homogeneous space for the algebraic torus  $R_{K/\mathbb{Q}}^1$ , and so the conjecture follows from work of Colliot-Thélène and Sansuc [6]. When  $r = 2$ , Heath-Brown and Skorobogatov [23] prove it under the additional assumption that  $\gcd(n, m_1, m_2) = 1$ , while Colliot-Thélène, Harari and Skorobogatov [9, Thm. 3.1] establish it in general. Our primary result establishes the conjecture for any  $r \geq 1$ .

**Theorem 1.1.** *The Brauer–Manin obstruction is the only obstruction to the Hasse principle and weak approximation on  $X^c$ .*

By combining Theorem 1.1 with the Brauer group calculation in [9, Cor. 2.7] we obtain the following corollary.

**Corollary 1.2.** *Suppose that  $\gcd(m_1, \dots, m_r) = 1$  and  $K$  does not contain a proper cyclic extension of  $\mathbb{Q}$ . Then  $X^c(\mathbb{Q}) \neq \emptyset$  and  $X^c$  satisfies weak approximation.*

In [23], for the first time, Heath-Brown and Skorobogatov combined the descent theory of Colliot-Thélène and Sansuc [7] with the Hardy–Littlewood circle method, in order to study the Hasse principle and weak approximation. In joint work with Skorobogatov [2], we introduced additive combinatorics into this subject and showed how it may usefully be combined with descent. This approach allowed us to study the variety  $X$  when  $K/\mathbb{Q}$  is quadratic. The case  $n = 2$  of Theorem 1.1 is a special case of [2, Thm 1.1]. Subsequently, Harpaz, Skorobogatov and Wittenberg [22] succeeded in showing how the finite complexity case of the generalised Hardy–Littlewood conjecture for primes, as established by Green and Tao [18] and Green–Tao–Ziegler [21], can be used in place of Schinzel’s hypothesis to study rational points on varieties using fibration arguments. Their work [22, Cor. 4.1] leads to a version of Theorem 1.1 in which the extension  $K/\mathbb{Q}$  is assumed to be cyclic, a fact that was previously only available under Schinzel’s hypothesis, as a special case of work by Colliot-Thélène and Swinnerton-Dyer [8] on pencils of Severi–Brauer varieties. Building on work of Wei [35], they also handle (see [22, Thm. 4.6]) the case in which  $K$  is a non-cyclic extension of  $\mathbb{Q}$  of prime degree such that the Galois group of the normal closure of  $K$  over  $\mathbb{Q}$  has a non-trivial abelian quotient. We emphasise that the results of the present paper are unconditional and make no assumptions on the degree of the field extension, nor upon the type of the extension, other than that the ground field is  $\mathbb{Q}$ .

Our approach is based upon the strategy of [2]. We use descent theory to reduce Theorem 1.1 to establishing the Hasse principle and weak approximation for some auxiliary varieties, which can be analysed using additive combinatorics. To introduce these varieties, let

$$f_1, \dots, f_r \in \mathbb{Q}[u_1, \dots, u_s]$$

be a system of pairwise non-proportional homogeneous linear polynomials, with  $s \geq 2$ . For each  $1 \leq i \leq r$ , let  $K_i$  denote a number field of degree  $n_i = [K_i : \mathbb{Q}] \geq 2$ . Central to our investigation will be the smooth variety  $\mathcal{V} \subset \mathbb{A}_{\mathbb{Q}}^{n_1 + \dots + n_r + s}$ , defined by

$$0 \neq \mathbf{N}_{K_i}(\mathbf{x}_i) = f_i(u_1, \dots, u_s), \quad (1 \leq i \leq r), \quad (1.3)$$

where  $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,n_i})$ . For this variety we establish the following theorem, whose proof forms the bulk of this paper.

**Theorem 1.3.** *The variety  $\mathcal{V}$  defined by (1.3) satisfies the Hasse principle and weak approximation.*

In fact (see Theorem 5.2) we shall produce an asymptotic formula for the number of suitably constrained integral points on  $\mathcal{V}$  of bounded height. When  $\mathcal{V}$  only involves quadratic extensions, Theorem 1.3 recovers [2, Thm. 1.2]. The latter result was established

using work of the second author [27, 28]. We will build on this work in order to obtain the general case of Theorem 1.3. When  $K_1, \dots, K_r$  are all assumed to be cyclic extensions of  $\mathbb{Q}$ , a shorter proof of Theorem 1.3 can be found in [22, Thm. 1.3].

**1.1. Overview.** We indicate how Theorem 1.3 implies Theorem 1.1 at the end of this introduction. The remainder of this paper is organised as follows. The overall goal is to prove Theorem 1.3 by asymptotically counting points of bounded height in  $\mathcal{V}(\mathbb{Z})$ , taking into account the additional constraints that are imposed by the weak approximation conditions. The associated counting function is introduced in Section 5. The asymptotic formula obtained in Theorem 5.2 for this counting function may prove to be of independent interest. Theorem 5.2 is proved using Green and Tao’s nilpotent Hardy–Littlewood method (see [18]) in combination with the Green–Tao–Ziegler inverse theorem [21].

While containing mostly classical material, Section 2 fixes the notation for the rest of the paper and describes a certain fundamental domain that is specific to our counting problems. Section 3 contains a variety of technical results required at later stages in the paper and may be consulted as needed. Section 4 studies the number of solutions to a congruence  $\mathbf{N}_K(\mathbf{x}) \equiv A \pmod{p^m}$ . These results are used in Section 5 in order to analyse the non-archimedean local densities that appear in the statement of Theorem 5.2. Section 6 establishes those estimates for the Green–Tao method that correspond to the minor arc estimates in the classical Hardy–Littlewood method. These are the estimates needed in order to apply the Green–Tao–Ziegler inverse theorem [21]. Section 7 generalises the construction of the divisor function majorant from [26] to a fairly wide class of positive multiplicative functions. Section 8 combines this majorant for a specific function with a sieve majorant (as appears in Green and Tao’s work [18] on primes) to form a majorant for our main counting function. Section 9 shows that this majorant is pseudorandom, which finally allows us in Section 10 to employ the Green–Tao method in combination with the inverse result [21] to prove Theorem 5.2.

**1.2. Descent.** We close our introduction with the deduction of Theorem 1.1 from Theorem 1.3. We use the construction of “vertical” torsors due to Schindler and Skorobogatov [31]. These are introduced in [31], in order to study varieties given by equations of a form similar to (1.3) via the circle method. Let  $\pi : X \rightarrow \mathbb{A}^1$  be the morphism which maps  $(t; \mathbf{x})$  to  $t$ . Let  $U_0 \subset \mathbb{A}^1$  be the open subset on which  $\prod_{i=1}^r (t - e_i) \neq 0$  and let  $U = \pi^{-1}(U_0)$ . Let  $T = R_{K/\mathbb{Q}}^1$  be the torus given by the affine equation  $\mathbf{N}_K(\mathbf{x}) = 1$ . In [31, §2], a partial compactification  $Y$  of  $X$  is constructed and vertical torsors  $\mathcal{T} \rightarrow Y$  are shown to exist. These are torsors  $\mathcal{T} \rightarrow Y$  whose type is the injective map of  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ -modules  $\hat{T}^r \rightarrow \text{Pic}(\overline{Y})$ . It follows from [31, Lemma 2.2] that the restriction  $\mathcal{T}_U$  of  $\mathcal{T}$  to  $U \subset Y$  is  $E \times V$ , where  $E$  is a principal homogeneous space for  $T$  and  $V \subset \mathbb{A}^{nr+1}$  is defined by

$$t - e_i = \lambda_i \mathbf{N}_K(\mathbf{x}_i) \neq 0, \quad (1 \leq i \leq r),$$

for  $\lambda_1, \dots, \lambda_r \in \mathbb{Q}^*$ . Finally, it follows from [31, Thm. 2.1] that Theorem 1.1 holds when  $V$  is shown to satisfy the Hasse principle and weak approximation for any  $\lambda_1, \dots, \lambda_r \in \mathbb{Q}^*$ .

But  $V$  is isomorphic to the variety cut out by the system of equations

$$e_1 - e_i = \lambda_i \mathbf{N}_K(\mathbf{x}_i) - \lambda_1 \mathbf{N}_K(\mathbf{x}_1), \quad (2 \leq i \leq r).$$

By an obvious change of variables it suffices to establish the Hasse principle and weak approximation for the variety in  $\mathbb{A}^{nr+r}$  defined by the system of equations

$$0 \neq (e_1 - e_i) \mathbf{N}_K(\mathbf{y}) = \lambda_i \mathbf{N}_K(\mathbf{x}_i) - \lambda_1 \mathbf{N}_K(\mathbf{x}_1), \quad (2 \leq i \leq r).$$

But this variety is isomorphic to the variety

$$\begin{aligned} v &= \mathbf{N}_K(\mathbf{y}) \neq 0, \\ u - e_i v &= \lambda_i \mathbf{N}_K(\mathbf{x}_i) \neq 0, \quad (1 \leq i \leq r), \end{aligned}$$

which is a special case of the varieties  $\mathcal{V}$  considered in Theorem 1.3. This concludes our deduction of Theorem 1.1 from Theorem 1.3.

**Notation.** In addition to the usual asymptotic notations, we write  $U \asymp V$  to mean that  $U \ll V$  and  $U \gg V$ , and we write  $V = U^{o(1)}$  to express that  $V = O_\delta(U^\delta)$  for every  $\delta > 0$ . If  $\mathcal{U}$  is a finite set, then we define  $\mathbb{E}_{u \in \mathcal{U}} = |\mathcal{U}|^{-1} \sum_{u \in \mathcal{U}}$ . We will write  $\mathbf{1}_{u \in \mathcal{P}}$ , or equivalently  $\mathbf{1}_{\mathcal{P}}(u)$ , to denote the characteristic function of an element  $u$  satisfying property  $\mathcal{P}$ .

**Acknowledgements.** While working on this paper the first author was supported by ERC grant 306457 and the second author was supported by EPSRC grant EP/E053262/1 and by ERC grant 208091. Some of this work was carried out during the programme “Arithmetic and geometry” in 2013 at the *Hausdorff Institute* in Bonn. We are grateful to J.-L. Colliot-Thélène and A. Skorobogatov for their interest in this work. We would also like to thank U. Derenthal for useful comments on Section 5 and A. Skorobogatov for pointing out a simplification in the descent argument above. Special thanks are due to the anonymous referee for numerous useful comments and for giving us a much simpler proof of Lemma 4.2 and Lemma 8.4.

## 2. ALGEBRAIC NUMBER THEORY

The purpose of this section is threefold. First, in Section 2.1, we recall mostly standard material from algebraic number theory (as found in [24] and [25]), in order to fix the notation for the rest of the paper. Next, in Section 2.2, we will turn to our specific situation. We will ultimately require a counting function that assigns to each integer  $m$  its number of representations by the norm form  $N_{K/\mathbb{Q}}(x_1\omega_1 + \cdots + x_n\omega_n)$ . Writing  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ , this problem will be turned into a finite counting problem by identifying representations  $m = N_{K/\mathbb{Q}}(\mathbf{x} \cdot \boldsymbol{\omega})$  and  $m = N_{K/\mathbb{Q}}(\mathbf{y} \cdot \boldsymbol{\omega})$  if  $\mathbf{x} \cdot \boldsymbol{\omega}$  and  $\mathbf{y} \cdot \boldsymbol{\omega}$  are associated by a unit in the ring of integers  $\mathfrak{o}_K$  of  $K$ . Such a unit is necessarily of norm  $+1$ . With this in mind, we will need to describe a fundamental domain for the action by (the free part of) the group of norm  $+1$  units, and its properties relevant to us. In particular, in order to apply a lattice point counting result from the geometry of numbers we will need to show that the regions we work with have a sufficiently nice boundary. Finally, Section 2.3 collects together some analytic information about the Dedekind zeta function.

**2.1. Three ways to view a number field.** Let  $K$  be a number field of degree  $n$  over  $\mathbb{Q}$ . We let  $D_K$  denote its discriminant, let  $\mathfrak{o} = \mathfrak{o}_K$  be the ring of integers and let  $U_K$  be the unit group. Given any  $\alpha \in K$  we will denote its norm by  $N_{K/\mathbb{Q}}(\alpha)$ . For any integral ideal  $\mathfrak{a} \subset \mathfrak{o}$  we write  $N\mathfrak{a} = \#\mathfrak{o}/\mathfrak{a}$  for its ideal norm.

Let  $r_1$  (resp.  $2r_2$ ) be the number of distinct real (resp. complex) embeddings of  $K$ . Hence  $n = r_1 + 2r_2$ . The  $r_1$  distinct real embeddings are denoted by  $\sigma_1, \dots, \sigma_{r_1}$ , while  $\sigma_{r_1+1}, \dots, \sigma_{r_1+2r_2}$  denote a complete set of  $2r_2$  distinct complex embeddings, with  $\sigma_{r_1+i}$  conjugate to  $\sigma_{r_1+r_2+i}$  for  $1 \leq i \leq r_2$ .

The map  $\varphi : \alpha \mapsto (\sigma_1(\alpha), \dots, \sigma_{r_1+r_2}(\alpha))$  canonically embeds  $K$  into the  $n$ -dimensional commutative  $\mathbb{R}$ -algebra  $V = K \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ . We will typically write  $v^{(l)}$ , for superscripts  $1 \leq l \leq r_1 + r_2$ , for the projection of any  $v \in V$  onto the  $l$ th component, which is uniquely determined by  $\varphi$ . Thus any  $v \in V$  can be written  $v = (v^{(1)}, \dots, v^{(r_1+r_2)})$ . We identify  $K$  with its image  $\varphi(K)$  in  $V$ . Under this identification our fixed  $\mathbb{Q}$ -basis  $\{\omega_1, \dots, \omega_n\}$  for  $K$  gives rise to an  $\mathbb{R}$ -basis  $\{\varphi(\omega_1), \dots, \varphi(\omega_n)\}$  for  $V$  and we may consider  $V$  to be the set  $\{x_1\omega_1 + \dots + x_n\omega_n : x_i \in \mathbb{R}\}$ . This allows us to associate to  $v \in V$  the corresponding vector  $\mathbf{x} \in \mathbb{R}^n$  and vice versa. For  $v \in V$  we define

$$\mathrm{Nm}(v) = v^{(1)} \dots v^{(r_1)} |v^{(r_1+1)}|^2 \dots |v^{(r_1+r_2)}|^2,$$

which in accordance with our convention, we shall also denote as  $\mathrm{Nm}(\mathbf{x})$ . This gives us a formal extension of the norm form  $\mathbf{N}_K : \mathbb{Q}^n \rightarrow \mathbb{Q}$  to  $\mathbb{R}^n$ . Indeed, if  $v = \varphi(\alpha)$  for some  $\alpha = x_1\omega_1 + \dots + x_n\omega_n \in K$ , then

$$\mathrm{Nm}(\varphi(\alpha)) = N_{K/\mathbb{Q}}(\alpha) = \mathrm{Nm}(\mathbf{x}) = \mathbf{N}_K(\mathbf{x}).$$

A third way of viewing  $K$  is through logarithmic coordinates (see [25, §5] for details). Writing  $V^* = (K \otimes_{\mathbb{Q}} \mathbb{R})^* \cong (\mathbb{R}^*)^{r_1} \times (\mathbb{C}^*)^{r_2}$ , we define the homomorphism  $L : V^* \rightarrow \mathbb{R}^{r_1+r_2}$  to be

$$v \mapsto (\log |v^{(1)}|, \dots, \log |v^{(r_1)}|, 2\log |v^{(r_1+1)}|, \dots, 2\log |v^{(r_1+r_2)}|).$$

Composing  $L$  with the embedding  $\varphi$  from above, we obtain the diagram

$$\begin{array}{ccc} K^* & \xrightarrow{\varphi} & V^* \\ & \searrow \psi & \downarrow L \\ & & \mathbb{R}^{r_1+r_2} \end{array}$$

where  $\psi = L \circ \varphi$ . For any  $\alpha \in K^*$ , the coordinate sum of  $\psi(\alpha)$  is given by

$$\psi(\alpha) \cdot \mathbf{1} = \log |N_{K/\mathbb{Q}}(\alpha)|,$$

where  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{r_1+r_2}$ . In particular  $\psi(U_K)$  is contained in the hyperplane

$$H = \{\mathbf{v} \in \mathbb{R}^{r_1+r_2} : \mathbf{v} \cdot \mathbf{1} = 0\} \subset \mathbb{R}^{r_1+r_2}. \quad (2.1)$$

If  $\mathbf{u}_{r_1+r_2} = (1, \dots, 1, 2, \dots, 2) \in \mathbb{R}^{r_1+r_2}$ , then  $\mathbf{u}_{r_1+r_2} \notin H$ . For any  $v \in V^*$  we may write

$$L(v) = \mathbf{u}_H + \xi_{r_1+r_2} \mathbf{u}_{r_1+r_2}$$

for some  $\mathbf{u}_H \in H$  and  $\xi_{r_1+r_2} \in \mathbb{R}$ . This decomposition allows us to understand easily the norm of an element  $v$ , since

$$\log |\mathrm{Nm}(v)| = L(v) \cdot \mathbf{1} = \xi_{r_1+r_2} \mathbf{u}_{r_1+r_2} \cdot \mathbf{1} = n \xi_{r_1+r_2}. \quad (2.2)$$

It follows that  $|\mathrm{Nm}(v)| \leq 1$  if and only if  $\xi_{r_1+r_2} \leq 0$ . Finally, note that  $\ker(\psi) = \mu_K$ , where  $\mu_K < U_K$  denotes the subgroup of roots of unity. Thus the map  $L$  separates the free part of the group of unity from its torsion part.

**2.2. Units of norm +1.** We are now ready to discuss the subgroup of  $U_K$  relevant to us and its action on  $V$ . Recall that  $N_{K/\mathbb{Q}}(\eta) = \pm 1$  for any  $\eta \in U_K$ . We shall work with the subgroup  $U_K^{(+)}$  of  $\eta \in U_K$  such that  $N_{K/\mathbb{Q}}(\eta) = 1$ . Since  $U_K^{(+)}$  is the kernel of the group homomorphism  $N_{K/\mathbb{Q}} : U_K \rightarrow \{\pm 1\}$ , we deduce that either  $U_K = U_K^{(+)}$  or  $U_K/U_K^{(+)} \cong \mathbb{Z}/2\mathbb{Z}$ . In particular,  $U_K$  and  $U_K^{(+)}$  share the same rank  $r = r_1 + r_2 - 1$ .

Let  $\mu_K^{(+)}$  be the norm +1 subgroup of  $\mu_K$ . Then we have decompositions

$$U_K = \mu_K \times Y_K, \quad U_K^{(+)} = \mu_K^{(+)} \times Y_K^{(+)},$$

where  $Y_K \cong Y_K^{(+)} \cong \mathbb{Z}^r$ . The relation between these decompositions can be described more precisely. This is only interesting in the case where  $U_K^{(+)}$  is a proper subgroup of  $U_K$ , which we assume for now. If  $K$  has a root of unity of norm  $-1$  then one can ensure that each generator of  $Y_K$  has norm 1, since each generator may be replaced by the product of itself and a root of unity. This allows for decompositions where  $Y_K = Y_K^{(+)}$ . If  $K$  has no root of unity of norm  $-1$ , then  $\mu_K = \mu_K^{(+)}$ , and one can ensure that exactly one generator of  $Y_K$  has norm  $-1$ . To see this, suppose  $\eta_1 \dots, \eta_r$  is a system of fundamental units such that  $N_{K/\mathbb{Q}}(\eta_1) = -1$ . We keep  $\eta_1$ , but replace any other generator  $\eta_i$  of norm  $-1$  by the product  $\eta_1 \eta_i$ . The resulting system of units  $\eta_1, \eta'_2 \dots, \eta'_r$ , say, still generates  $Y_K$  and has the required property. Furthermore,  $\eta_1^2, \eta'_2 \dots, \eta'_r$  forms a system of generators for  $Y_K^{(+)}$ .

Our main interest in  $Y_K^{(+)}$  lies in the action it induces on  $V$  and the associated coordinate space  $\mathbb{R}^n$ . In general, the action of  $U_K$  on  $V$  by multiplication induces a natural action on the coordinate space  $\mathbb{R}^n$  as follows. For any pair  $(\eta, \mathbf{x}) \in U_K \times \mathbb{R}^n$  we let  $\eta \cdot \mathbf{x} = \mathbf{y} \in \mathbb{R}^n$  denote the coordinate vector of  $\eta(x_1 \omega_1 + \dots + x_n \omega_n) = y_1 \omega_1 + \dots + y_n \omega_n$ .

We require a fundamental domain  $\mathfrak{F}_+ \subset V^*$  for the action of  $\varphi(Y_K^{(+)})$  on  $V^*$  that is explicit enough to allow lattice point counting arguments to be applied. In the case where  $Y_K^{(+)}$  is replaced by  $Y_K$ , the construction of such a domain is classical (see [25, §§5,6]), and it is not difficult to adapt the construction so as to apply to our situation. This construction builds on the observation that the action of  $U_K$  is easier to understand in the logarithmic space. This is useful since the restriction of  $L$  to  $\varphi(Y_K^{(+)})$  is an isomorphism, which allows us to describe a fundamental domain for the action of  $\varphi(Y_K^{(+)})$  on  $V^*$  in terms of a fundamental domain for the action of  $\psi(Y_K^{(+)})$  on  $\mathbb{R}^{r_1+r_2}$ .

Let  $\delta_1, \dots, \delta_r$  be generators for  $Y_K^{(+)}$  and let  $\mathbf{u}_i = \psi(\delta_i)$  for  $1 \leq i < r_1 + r_2$ . Then  $\psi(U_K^{(+)}) = \psi(Y_K^{(+)})$  is a lattice of rank  $r$  contained in the hyperplane  $H$  that was defined in (2.1). We denote this lattice by  $\Lambda_+$  and note that it is generated by  $\mathbf{u}_1, \dots, \mathbf{u}_{r_1+r_2-1}$ .



Let  $F_+ \subset H$  be a fundamental parallelootope for  $\Lambda_+$ , and recall that the vector  $\mathbf{u}_{r_1+r_2} = (1, \dots, 1, 2, \dots, 2) \in \mathbb{R}^{r_1+r_2}$  does not belong to  $H$ . Thus  $F_+ \oplus \mathbb{R}\mathbf{u}_{r_1+r_2}$  describes a fundamental domain for the action of  $\psi(Y_K^{(+)})$  on  $\psi(K^*) = \mathbb{R}^{r_1+r_2}$ . Since  $Y_K^{(+)}$  maps isomorphically onto  $\Lambda_+$ , an application of [25, Lemma 1 in §6] implies the following result.

**Lemma 2.1.** *The set  $\mathfrak{F}_+ = \{v \in V^* : L(v) \in F_+ \oplus \mathbb{R}\mathbf{u}_{r_1+r_2}\}$  is a fundamental domain for the action of  $\varphi(Y_K^{(+)})$  on  $V^*$ .*

We now turn to the desirable properties of the domain  $\mathfrak{F}_+$ , that ultimately facilitate lattice point counting. Recall that a region  $S \subset \mathbb{R}^n$  is said to be a *cone* when  $\mathbf{x} \in S$  if and only if  $\lambda\mathbf{x} \in S$ , for any  $\lambda \in \mathbb{R}_{>0}$ . Moreover, if  $S \subset \mathbb{R}^n$  is bounded, its boundary is called  *$(n-1)$ -Lipschitz parametrisable* (see [25, p.166]) if it is contained in the union of the images of finitely many Lipschitz functions  $f : [0, 1]^{n-1} \rightarrow \mathbb{R}^n$ . It is easy to see that  $\mathfrak{F}_+$  is a cone. We are interested in the set

$$\begin{aligned} \mathfrak{F}_+(1) &= \{v \in \mathfrak{F}_+ : |\mathrm{Nm}(v)| \leq 1\} \\ &= \{v \in V^* : L(v) \in F_+ \oplus \mathbb{R}_{\leq 0}\mathbf{u}_{r_1+r_2}\}, \end{aligned}$$

where the second equality follows from (2.2). The proof contained in [25, pp.168–172] applies *mutatis mutandis* to our situation and establishes the following result.

**Lemma 2.2.** *The domain  $\mathfrak{F}_+(1)$  has an  $(n-1)$ -Lipschitz parametrisable boundary.*

We will mainly be working in the coordinate space  $\mathbb{R}^n$ . The map

$$v : \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow (\mathbb{R}^*)^{r_1} \times (\mathbb{C}^*)^{r_2}$$

that takes  $\mathbf{x}$  to  $v(\mathbf{x}) = x_1\omega_1 + \dots + x_n\omega_n = (v^{(1)}, \dots, v^{(r_1+r_2)})$  is a linear isomorphism and preserves Lipschitz parametrisability. In particular, if

$$\mathfrak{D}_+ = \{\mathbf{x} \in \mathbb{R}^n : v(\mathbf{x}) \in \mathfrak{F}_+\} \quad (2.3)$$

denotes the preimage of the fundamental domain in  $\mathbb{R}^n$ , and if

$$\mathfrak{D}_+(1) = \{\mathbf{x} \in \mathfrak{D}_+ : |\mathbf{N}_K(\mathbf{x})| \leq 1\},$$

then Lemma 2.2 implies that  $\mathfrak{D}_+(1)$  has an  $(n-1)$ -Lipschitz parametrisable boundary.

We slightly refine the sets under consideration. The sign of  $\mathbf{N}_K$  is invariant under the action of  $Y_K^{(+)}$ . Thus, for  $\epsilon \in \{\pm\}$  and  $T > 0$  we define the sets

$$\mathfrak{D}_+^\epsilon = \{\mathbf{x} \in \mathfrak{D}_+ : 0 < \epsilon \mathbf{N}_K(\mathbf{x})\}$$

and

$$\mathfrak{D}_+^\epsilon(T) = \{\mathbf{x} \in \mathfrak{D}_+ : 0 < \epsilon \mathbf{N}_K(\mathbf{x}) \leq T\}. \quad (2.4)$$

Since  $\mathfrak{F}_+$  is a cone, the same is true for  $\mathfrak{D}_+$  and  $\mathfrak{D}_+^\epsilon$ . We deduce that  $\mathfrak{D}_+^\epsilon(1)$  has an  $(n-1)$ -Lipschitz parametrisable boundary from the same property for  $\mathfrak{D}_+(1)$ . Furthermore, we have  $\mathfrak{D}_+^\epsilon(T) = T^{1/n}\mathfrak{D}_+^\epsilon(1)$ .

Note that the same facts hold true in the classical setting for

$$\mathfrak{D}^\epsilon(T) = \{\mathbf{x} \in \mathfrak{D} : 0 < \epsilon \mathbf{N}_K(\mathbf{x}) \leq T\}, \quad (2.5)$$

where  $\mathfrak{D} = \{\mathbf{x} \in \mathbb{R}^n : v(\mathbf{x}) \in \mathfrak{F}\}$  and  $\mathfrak{F} \subset V^*$  is the fundamental domain for  $\varphi(Y_K)$ .

**2.3. Dirichlet coefficients of  $\zeta_K$ .** The construction of the majorant in Section 8 relies on a careful analysis of the sequence of Dirichlet coefficients of the Dedekind zeta function of a number field  $K$ . Here we recall the essential properties of  $\zeta_K$  and its Dirichlet coefficients, as found in Landau [24] or Marcus [25], and deduce some preliminary facts required in Section 8.

The Dedekind zeta function is defined to be

$$\zeta_K(s) = \sum_{(0) \neq \mathfrak{a} \subset \mathfrak{o}} \frac{1}{(\mathbf{N} \mathfrak{a})^s} = \sum_{m=1}^{\infty} \frac{r_K(m)}{m^s}, \quad (2.6)$$

for  $s \in \mathbb{C}$  with  $\Re(s) > 1$ , with

$$r_K(m) = \#\{\mathfrak{a} \subset \mathfrak{o} : \mathbf{N} \mathfrak{a} = m\}.$$

The Dedekind zeta function admits a meromorphic continuation to all of  $\mathbb{C}$  with a simple pole at  $s = 1$  and  $\text{Residue}_{s=1} \zeta_K(s) = h\kappa$ , where  $h$  is the class number,

$$\kappa = \frac{2^{r_1} (2\pi)^{r_2} R_K}{|\mu_K| \sqrt{|D_K|}}, \quad (2.7)$$

and  $R_K$  is the regulator. It follows from [25, Thms. 39 and 40] that

$$\sum_{m \leq x} r_K(m) = h\kappa x + O(x^{1-1/n}), \quad (2.8)$$

so that the average order of  $r_K$  is constant.

The function  $r_K$  is multiplicative. To describe its behaviour at prime powers, let  $p$  be any rational prime and recall that the principal ideal  $(p)$  factorises into a product of prime ideals in  $\mathfrak{o}$ . That is,

$$(p) = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}, \quad (2.9)$$

where  $e_i = e_{\mathfrak{p}_i}(p)$ ,  $r = r(p) \in \mathbb{Z}_{>0}$  and each  $\mathfrak{p}_i \subset \mathfrak{o}$  is a prime ideal satisfying  $\mathbf{N} \mathfrak{p}_i = p^{f_i}$ , for some  $f_i = f_{\mathfrak{p}_i}(p) \in \mathbb{Z}_{>0}$ . As in [25, §3], we have  $\sum_{i=1}^r e_i f_i = n$ . Thus

$$r_K(p^m) = \#\{\mathfrak{p}_1^{m_1} \cdots \mathfrak{p}_r^{m_r} \subset \mathfrak{o} : f_1 m_1 + \cdots + f_r m_r = m\},$$

for any  $m \in \mathbb{Z}_{>0}$ . It follows from this that

$$r_K(p^m) \leq (m+1)^n. \quad (2.10)$$

At rational primes we obtain

$$r_K(p) = \#\{i \in \{1, \dots, r\} : f_i = 1\} = \#\{\mathfrak{p} \mid (p) : f_{\mathfrak{p}}(p) = 1\}. \quad (2.11)$$

In view of (2.11), we partition the set of rational primes into three sets

$$\begin{aligned} \mathscr{P}_0 &= \{p \mid D_K\}, \\ \mathscr{P}_1 &= \{p \nmid D_K : \exists \mathfrak{p} \mid (p) \text{ such that } f_{\mathfrak{p}}(p) = 1\}, \\ \mathscr{P}_2 &= \{p \nmid D_K : f_{\mathfrak{p}}(p) \geq 2 \ \forall \mathfrak{p} \mid (p)\}. \end{aligned} \quad (2.12)$$

The contributions to  $r_K$  from  $\mathscr{P}_0 \cup \mathscr{P}_1$  and from  $\mathscr{P}_2$  will be dealt with separately.

We end this section with some technical results concerning the restricted Euler product

$$F(s) = \prod_{p \in \mathcal{P}_2} \left(1 - \frac{1}{p^s}\right)^{-1}, \quad (2.13)$$

for  $s \in \mathbb{C}$  with  $\Re(s) > 1$ . The following result describes the analytic structure of  $F(s)$ .

**Lemma 2.3.** *There exists  $\delta \in \mathbb{Q}$  satisfying  $1/n \leq \delta \leq 1$  such that  $\mathcal{P}_1$  has Dirichlet density  $\delta$ . Furthermore, there is a function  $G(s)$ , which is holomorphic and non-zero in the closed half-plane  $\Re(s) \geq 1$ , such that  $F(s) = \zeta(s)^{1-\delta} G(s)$ .*

*Proof.* The first part follows from the Čebotarev density theorem (cf. [30, Cor 13.6]), with  $\delta = 1/n$  if and only if  $K/\mathbb{Q}$  is a Galois extension. This implies that there exists a function  $G_1(s)$ , which is holomorphic and non-zero in the closed half-plane  $\Re(s) \geq 1$ , such that

$$\prod_{p \in \mathcal{P}_1} \left(1 - \frac{1}{p^s}\right)^{-1} = \zeta(s)^\delta G_1(s). \quad (2.14)$$

On the other hand,

$$\prod_{p \in \mathcal{P}_1} \left(1 - \frac{1}{p^s}\right)^{-1} = \zeta(s) \prod_{p \in \mathcal{P}_0 \cup \mathcal{P}_2} \left(1 - \frac{1}{p^s}\right) = \zeta(s) F(s)^{-1} G_2(s),$$

where  $G_2(s)$  is entire and non-zero. Combining these expressions we conclude the proof of the lemma by taking  $G(s) = G_1(s)^{-1} G_2(s)$ .  $\square$

**Corollary 2.4.** *We have*

$$\prod_{\substack{p \in \mathcal{P}_2 \\ p \leq T}} \left(1 - \frac{1}{p}\right)^{-1} \asymp (\log T)^{1-\delta}.$$

*Proof.* By means of the Tauberian theorem [29, Thm. 5.11], applied to the Dirichlet series  $F(s)$ , we deduce from Lemma 2.3 that

$$(\log T)^{1-\delta} \asymp \sum_{m \leq T} \frac{\mathbf{1}_{\langle \mathcal{P}_2 \rangle}(m)}{m} \leq \prod_{\substack{p \in \mathcal{P}_2 \\ p \leq T}} \left(1 - \frac{1}{p}\right)^{-1}.$$

This gives the correct lower bound. To establish the upper bound, we deduce from (2.14) that

$$(\log T)^\delta \asymp \sum_{m \leq T} \frac{\mathbf{1}_{\langle \mathcal{P}_1 \rangle}(m)}{m} \leq \prod_{\substack{p \in \mathcal{P}_1 \\ p \leq T}} \left(1 - \frac{1}{p}\right)^{-1}.$$

But then

$$(\log T)^\delta \prod_{\substack{p \in \mathcal{P}_2 \\ p \leq T}} \left(1 - \frac{1}{p}\right)^{-1} \ll \prod_{\substack{p \in \mathcal{P}_1 \cup \mathcal{P}_2 \\ p \leq T}} \left(1 - \frac{1}{p}\right)^{-1} \ll \log T,$$

as required.  $\square$

### 3. TECHNICAL TOOLS

**3.1. Geometry of numbers.** We will need to be able to estimate the number of lattice points in shifts of sufficiently well-behaved expanding regions. Let  $n \in \mathbb{Z}_{>0}$  and let  $\mathcal{B}$  be any bounded subset of  $\mathbb{R}^n$ . Write  $T\mathcal{B} = \{T\mathbf{x} : \mathbf{x} \in \mathcal{B}\}$  for the dilation by  $T > 0$ . The following result is classical.

**Lemma 3.1.** *Assume that  $\mathcal{B}$  is bounded and that for any  $\varepsilon \in (0, 1)$  the  $\varepsilon$ -neighbourhood of the boundary  $\partial\mathcal{B}$  has volume  $O(\varepsilon)$ . Let  $\mathbf{a} \in \mathbb{R}^n$  and let  $T \geq 1$ . Then*

$$\#(\mathbb{Z}^n \cap (T\mathcal{B} + \mathbf{a})) = \text{vol}(\mathcal{B})T^n + O(T^{n-1}).$$

*Proof.* Observe that the  $\varepsilon$ -neighbourhood of  $\partial T\mathcal{B}$  arises as dilation by  $T$  of the  $\varepsilon T^{-1}$ -neighbourhood of  $\partial\mathcal{B}$  and has volume  $O(\varepsilon T^{n-1})$ . The lemma follows (cf. [18, App. A]) since

$$\#(\mathbb{Z}^n \cap (T\mathcal{B} + \mathbf{a})) = \text{vol}((\mathbb{Z}^n \cap (T\mathcal{B} + \mathbf{a})) + [0, 1]^n),$$

and the set in the latter volume agrees with  $T\mathcal{B} + \mathbf{a}$  outside an  $O(1)$ -neighbourhood of  $\partial T\mathcal{B} + \mathbf{a}$ .  $\square$

It is not hard to see that any non-empty bounded set in  $\mathbb{R}^n$  whose boundary is  $(n-1)$ -Lipschitz parametrisable satisfies the hypotheses of Lemma 3.1. This follows, for example, from the proof of Lemma 2 in [25, §6]. Similarly, bounded convex sets in  $\mathbb{R}^n$  satisfy the hypotheses of the lemma (see [18, Cor. A.2], for example).

Given a finite set of fixed regions  $\mathcal{B}_1, \dots, \mathcal{B}_m$  to which Lemma 3.1 applies, the hypotheses of the lemma are also met by any set which arises through unions and intersections of these sets. In particular it applies to intersections of bounded convex sets with bounded sets having  $(n-1)$ -Lipschitz parametrisable boundary.

**3.2. Complex analysis.** Throughout this section we will write  $\sigma$  for the real part of a complex number  $s \in \mathbb{C}$ . In the course of Sections 8 and 9 we will encounter several truncated Euler products of the following form. For a given constant  $C > 0$ , given  $x \geq 1$  and a given multiplicative arithmetic function  $h : \mathbb{Z}_{>0} \rightarrow \mathbb{C}$ , define the Euler product

$$E_{C,x}(s; h) = \prod_{C < p < x} \left( 1 + \sum_{k \geq 1} \frac{h(p^k)}{p^{sk}} \right),$$

for  $\sigma > 1$ . Since the product is truncated, one expects that  $E_{C,x}(1 + s_0; h)$  is well approximated by its value at  $s = 1$ , provided that  $|s_0|$  is sufficiently small and one has some control on  $h$ . The following result makes this statement precise.

**Lemma 3.2.** *Let  $c > 0$  be a constant. Let  $H \geq 1$  and suppose  $h : \mathbb{Z}_{>0} \rightarrow \mathbb{C}$  is a multiplicative function satisfying  $|h(p^k)| \leq H^k$  at all prime powers  $p^k$ . Then the Euler product  $E(s) = E_{3eH,x}(s; h)$  satisfies*

$$E(1 + s_0) = E(1) + O(|s_0|(\log x)^{O(1)}).$$

uniformly in  $x$ , for  $s_0 \in \mathbb{C}$  with  $|s_0| \leq x^{-c}$ . Furthermore, we have

$$|E(1)| \asymp \left| \prod_{3eH < p < x} \left( 1 + \frac{h(p)}{p} \right) \right|. \quad (3.1)$$

The implied constants in these estimates are allowed to depend on  $c$  and  $H$ .

For  $h$  satisfying  $|h(p)| \leq H$  on the primes, we may combine (3.1) with Mertens's theorem to deduce that

$$(\log x)^{-H} \ll |E(1)| \ll (\log x)^H.$$

This shows that the main term dominates the error term in our asymptotic formula for  $E(1+s)$ , when  $|s_0| \leq x^{-c}$ .

*Proof of Lemma 3.2.* We allow our implied constants to depend on  $c$  and  $H$ . Let  $E(s) = E_{3eH,x}(s; h)$  and let  $s \in \mathbb{C}$  be such that  $|s| < (\log x)^{-1}$ . Then

$$\left| \sum_{k \geq 1} \frac{h(p^k)}{p^{(1+s)k}} \right| \leq \frac{H}{p^{1+\sigma} - H}.$$

This is at most  $1/2$  for  $3eH < p < x$ , since

$$p^{1+\sigma} \geq p^{1-(\log x)^{-1}} \geq pe^{-1} \geq 3H.$$

This shows that  $E(1+s)$  is non-zero for  $s$  satisfying  $|s| < (\log x)^{-1}$  and, furthermore, that  $E$  is holomorphic on a domain containing this disc. The Taylor expansion about 1 is given by

$$E(1+s) = \sum_{j \geq 0} s^j \frac{E^{(j)}(1)}{j!}.$$

Cauchy's inequality yields

$$\frac{|E^{(j)}(1)|}{j!} \leq (\log x)^j \max_{|s|=(\log x)^{-1}} |E(1+s)|.$$

But the right hand side is bounded by

$$\begin{aligned} &\leq (\log x)^j \prod_{3eH < p < x} \left( 1 + \sum_{k \geq 1} \frac{|h(p^k)|}{p^{(1-1/\log x)k}} \right) \\ &\leq (\log x)^j \prod_{3eH < p < x} \left( 1 + \frac{H}{p^{1-\frac{1}{\log x}} - H} \right) \\ &\leq (\log x)^j \prod_{3eH < p < x} \left( 1 + \frac{3eH}{2p} \right) \\ &\ll (\log x)^{j+O(1)}. \end{aligned}$$

Thus for  $|s_0| < x^{-c}$  we have

$$|E(1) - E(1 + s_0)| \ll \sum_{j \geq 1} |s_0|^j (\log x)^{j+O(1)} \leq |s_0| (\log x)^{O(1)}.$$

To check the final claim of the lemma, we recall that  $|h(p^k)| \leq H^k$ . Using the logarithmic series we therefore deduce that

$$\begin{aligned} \log \left( E(1) \prod_{3eH < p < x} \left( 1 + \frac{h(p)}{p} \right)^{-1} \right) &= \sum_{3eH < p < x} \left( \log \left( 1 + \sum_{k \geq 1} \frac{h(p^k)}{p^k} \right) - \log \left( 1 + \frac{h(p)}{p} \right) \right) \\ &= \sum_{3eH < p < x} \left( \sum_{k \geq 2} \frac{h(p^k)}{p^k} + O\left(\frac{H^2}{p^2}\right) \right). \end{aligned}$$

But this is  $O(1)$ , which therefore concludes the proof.  $\square$

**3.3. Lifting lemmas.** This section establishes two fairly general results of Hensel type, the second of which will be applied in Sections 4, 5, 6 and 10. Let  $p$  denote a prime number and let  $v_p(n_1, \dots, n_s)$  denote the  $p$ -adic order of the greatest common divisor of any  $s$ -tuple of integers  $(n_1, \dots, n_s)$ .

**Lemma 3.3.** *Let  $m, \ell, \delta \in \mathbb{Z}_{\geq 0}$ , with*

$$m \geq 2\delta + 1, \quad 0 \leq \delta \leq m - \ell.$$

*Suppose we are given a polynomial  $F \in \mathbb{Z}[t_1, \dots, t_s]$ ,  $A \in \mathbb{Z}$  and  $\mathbf{a} \in \mathbb{Z}^s$ . Let*

$$R_\delta(p^m, A; p^\ell) = \left\{ \mathbf{t} \in (\mathbb{Z}/p^m\mathbb{Z})^s : \begin{array}{l} F(\mathbf{t}) \equiv A \pmod{p^m}, \quad v_p(\nabla F(\mathbf{t})) = \delta \\ \mathbf{t} \equiv \mathbf{a} \pmod{p^\ell} \end{array} \right\}.$$

*Then we have*

$$\frac{\#R_\delta(p^m, A; p^\ell)}{p^{m(s-1)}} = \frac{\#R_\delta(p^{m+1}, A + kp^m; p^\ell)}{p^{(m+1)(s-1)}},$$

*uniformly for  $k \in \mathbb{Z}/p\mathbb{Z}$ .*

*Proof.* For any  $\mathbf{t} \in R_\delta(p^m, A; p^\ell)$  and any  $\mathbf{t}' \in \mathbb{Z}^s$ , the condition  $m \geq 2\delta + 1$  implies that

$$\begin{aligned} F(\mathbf{t} + p^{m-\delta}\mathbf{t}') &\equiv F(\mathbf{t}) + p^{m-\delta}\mathbf{t}' \cdot \nabla F(\mathbf{t}) \pmod{p^m} \\ &\equiv A \pmod{p^m}. \end{aligned}$$

Similarly, we deduce that

$$\begin{aligned} \nabla F(\mathbf{t} + p^{m-\delta}\mathbf{t}') - \nabla F(\mathbf{t}) &\equiv \mathbf{0} \pmod{p^{m-\delta}} \\ &\equiv \mathbf{0} \pmod{p^{\delta+1}}, \end{aligned}$$

and, since  $\ell \leq m - \delta$ , we also have  $\mathbf{t} + p^{m-\delta}\mathbf{t}' \equiv \mathbf{a} \pmod{p^\ell}$ . Thus  $R_\delta(p^m, A; p^\ell)$  consists of cosets modulo  $p^{m-\delta}$ .

Let  $\mathbf{t} \in (\mathbb{Z}/p^{m+1}\mathbb{Z})^s$  such that  $\mathbf{t} \pmod{p^m} \in R_\delta(p^m, A; p^\ell)$ . Then  $\mathbf{t} + p^{m-\delta}\mathbf{t}'$  runs through  $p^s$  different cosets modulo  $p^{m+1-\delta}$  as  $\mathbf{t}'$  runs through  $\mathbb{Z}^s$ . Moreover, for any  $k \in \mathbb{Z}/p\mathbb{Z}$ , we have  $\mathbf{t} + p^{m-\delta}\mathbf{t}' \in R_\delta(p^{m+1}, A + kp^m; p^\ell)$  if and only if

$$p^{-m}(F(\mathbf{t}) - A) + p^{-\delta}\mathbf{t}' \cdot \nabla F(\mathbf{t}) \equiv k \pmod{p},$$

for which there are precisely  $p^{s-1}$  incongruent solutions in  $\mathbf{t}'$  modulo  $p$ . This establishes the lemma.  $\square$

Now let  $G \in \mathbb{Z}[x_1, \dots, x_n]$  be a homogeneous polynomial of degree  $n$  and let  $A \in \mathbb{Z}$ . For given  $\mathbf{a} \in \mathbb{Z}^n$  and  $m, \ell \in \mathbb{Z}_{\geq 0}$ , let

$$\gamma(p^m, A; p^\ell) = \# \left\{ \mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : \begin{array}{l} G(\mathbf{x}) \equiv A \pmod{p^m} \\ \mathbf{x} \equiv \mathbf{a} \pmod{p^\ell} \end{array} \right\}.$$

The counting function  $\gamma$  satisfies the following lifting property.

**Lemma 3.4** (cf. [27, Cor. 6.4]). *Assume that  $m \geq 1$ ,  $A \neq 0$  and*

$$\ell + v_p(A) + v_p(n) < \frac{m}{2}.$$

*Then we have*

$$\frac{\gamma(p^m, A; p^\ell)}{p^{m(n-1)}} = \frac{\gamma(p^{m+1}, A + kp^m; p^\ell)}{p^{(m+1)(n-1)}},$$

*uniformly for  $k \in \mathbb{Z}/p\mathbb{Z}$ .*

*Proof.* Let  $m \geq 1$  and let  $\mathbf{x} \pmod{p^m}$  be such that  $G(\mathbf{x}) \equiv A \pmod{p^m}$  and  $\mathbf{x} \equiv \mathbf{a} \pmod{p^\ell}$  and  $p^\delta \mid \nabla G(\mathbf{x})$ , for some  $\delta \geq 0$ . Since  $\mathbf{x} \cdot \nabla G(\mathbf{x}) = nG(\mathbf{x})$  we conclude that  $p^{\min\{\delta, m\}} \mid nA$ , whence  $\delta < \frac{m}{2} - \ell$  under the hypotheses of the lemma. In particular this inequality implies that  $m \geq 2\delta + 1$  and  $\delta \leq m - \ell$ . Observe that

$$\begin{aligned} \gamma(p^m, A; p^\ell) &= \sum_{0 \leq \delta < \frac{m}{2} - \ell} \# R_\delta(p^m, A; p^\ell), \\ \gamma(p^{m+1}, A + kp^m; p^\ell) &= \sum_{0 \leq \delta < \frac{m}{2} - \ell} \# R_\delta(p^{m+1}, A + kp^m; p^\ell), \end{aligned}$$

for any  $k \in \mathbb{Z}/p\mathbb{Z}$ , in the notation of Lemma 3.3. The statement of Lemma 3.4 therefore follows from Lemma 3.3 with  $F = G$  and  $s = n$ .  $\square$

#### 4. NORM FORMS MODULO $p^m$

Throughout this section  $K/\mathbb{Q}$  will denote a finite extension of degree  $n$ , with integral basis  $\{\omega_1, \dots, \omega_n\}$  for the ring of integers  $\mathfrak{o} = \mathfrak{o}_K$ . Suppose we are given an integral ideal  $\mathfrak{a} \subset \mathfrak{o}$ , with corresponding  $\mathbb{Z}$ -basis  $\{\alpha_1, \dots, \alpha_n\}$ . These bases are both  $\mathbb{Q}$ -bases for  $K/\mathbb{Q}$ . We let  $\Delta(\alpha_1, \dots, \alpha_n) = |\det(\sigma_i(\alpha_j))|^2$ , and similarly for  $\{\omega_1, \dots, \omega_n\}$ . Let  $c_{k\ell} \in \mathbb{Z}$  be such that

$$\alpha_k = \sum_{\ell=1}^n c_{k\ell} \omega_\ell, \tag{4.1}$$

for  $1 \leq k \leq n$ . Then according to [24, Satz 40 and 103], we have

$$\Delta(\alpha_1, \dots, \alpha_n) = (N\mathfrak{a})^2 |D_K| = |\det(c_{k\ell})|^2 \Delta(\omega_1, \dots, \omega_n).$$

In particular  $N\mathfrak{a} = |\det(c_{k\ell})|$ .

The norm forms we discuss in this section take the more general shape

$$N(\mathbf{x}; \mathfrak{a}) = N_{K/\mathbb{Q}}(x_1\alpha_1 + \dots + x_n\alpha_n), \quad (4.2)$$

which defines a homogeneous polynomial of degree  $n$  with coefficients in  $\mathbb{Z}$ . Note that  $N(\mathbf{x}; \mathfrak{o}) = N_K(\mathbf{x})$  in our earlier notation, which we will often abbreviate by  $N(\mathbf{x})$ . Given  $A \in \mathbb{Z}$ ,  $\mathbf{x}_0 \in \mathbb{Z}^n$  and  $M, q \in \mathbb{Z}_{>0}$  with  $M \mid q$ , we define the counting function

$$\varrho(q, A, \mathfrak{a}; M) = \# \left\{ \mathbf{x} \in (\mathbb{Z}/q\mathbb{Z})^n : \begin{array}{l} N(\mathbf{x}; \mathfrak{a}) \equiv A \pmod{q} \\ \mathbf{x} \equiv \mathbf{x}_0 \pmod{M} \end{array} \right\}. \quad (4.3)$$

Such counting functions appear naturally when analysing weak approximation conditions at non-archimedean places. In the special case  $\mathfrak{a} = \mathfrak{o}$ , we put

$$\varrho(q, A; M) = \varrho(q, A, \mathfrak{o}; M). \quad (4.4)$$

Likewise, when  $M = 1$ , we set

$$\varrho(q, A, \mathfrak{a}) = \varrho(q, A, \mathfrak{a}; 1), \quad \varrho(q, A) = \varrho(q, A, \mathfrak{o}; 1).$$

This section is devoted to a detailed analysis of the quantities  $\varrho(q, A, \mathfrak{a})$  and  $\varrho(q, A)$ . When  $M \neq 1$  it will suffice for our purposes to note that  $\varrho(q, A, \mathfrak{a}; M) \leq \varrho(q, A, \mathfrak{a})$  and apply the results for  $M = 1$ . By the Chinese remainder theorem we may consider  $\varrho(q, A, \mathfrak{a})$  and  $\varrho(q, A)$  in the special case  $q = p^m$  for a rational prime  $p$  and  $m \in \mathbb{Z}_{>0}$ . We will mainly be concerned with the situation for  $p \nmid D_K N\mathfrak{a}$ . Our first result shows that any two norm forms are locally equivalent.

**Lemma 4.1** (cf. [27, Lemma 4.2]). *Let  $m \in \mathbb{Z}_{>0}$  and let  $p \nmid N\mathfrak{a}$ . Then we have*

$$\varrho(p^m, A, \mathfrak{a}) = \varrho(p^m, A).$$

*Proof.* Let  $\mathbf{C} \in M_n(\mathbb{Z})$  be the matrix with coefficients  $c_{k\ell}$  as in (4.1). Since  $N\mathfrak{a} = |\det \mathbf{C}|$ , it follows that  $p \nmid \det \mathbf{C}$ , whence  $\mathbf{C}$  and  $\mathbf{C}^t$  are invertible in  $\mathbb{Z}_p$ . Let  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ . Then, for any  $\mathbf{x} \in \mathbb{Z}^n$ , we have

$$\begin{aligned} N(\mathbf{x}\mathbf{C}^t; \mathfrak{o}) &= N_{K/\mathbb{Q}}(\boldsymbol{\omega} \cdot (\mathbf{x}\mathbf{C}^t)) \\ &= N_{K/\mathbb{Q}}((\mathbf{C}\boldsymbol{\omega}) \cdot \mathbf{x}) \\ &= N(\mathbf{x}; \mathfrak{a}). \end{aligned}$$

It follows that  $N(\mathbf{x}; \mathfrak{a})$  and  $N(\mathbf{x}; \mathfrak{o})$  are equivalent over  $\mathbb{Z}_p$ , which suffices for the lemma.  $\square$

We now have everything in place to record our main result in this section.

**Lemma 4.2.** *Let  $A \in \mathbb{Z}$ , let  $m \in \mathbb{Z}_{>0}$ , let  $p$  be a prime and let  $k = v_p(A)$ . Then we have*

$$\frac{\varrho(p^m, A)}{p^{m(n-1)}} \ll \min\{k+1, m\}^n.$$



Suppose that  $p \nmid D_K$  and  $k < m$ . Then we have

$$\frac{\varrho(p^m, A)}{p^{m(n-1)}} = r_K(p^k) \left(1 - \frac{1}{p}\right)^{-1} \prod_{\mathfrak{p}|p} \left(1 - \frac{1}{N \mathfrak{p}}\right).$$

*Proof.* Our proof of this result was suggested to us by the anonymous referee and is based on the observation that  $\varrho(p^m, A)$  is equal to the number of  $\alpha \in \mathfrak{o}$ , modulo  $p^m$ , for which  $N_{K/\mathbb{Q}}(\alpha) \equiv A \pmod{p^m}$ . It will be convenient to temporarily abbreviate  $N_{K/\mathbb{Q}}$  by  $N$  in what follows.

We first consider the special case where  $p^m \mid A$ . In this case every  $\alpha$  that is counted by  $\varrho(p^m, A)$  will have an ideal divisor  $\mathfrak{q}$  such that  $p^m \mid N \mathfrak{q}$ , and with the property that  $p^m \nmid N \mathfrak{q}'$  for every proper divisor  $\mathfrak{q}' \mid \mathfrak{q}$ . Thus  $\mathfrak{q} = \prod_{\mathfrak{p}} \mathfrak{p}^{e_{\mathfrak{p}}}$  for prime ideal divisors  $\mathfrak{p} \mid (p)$ , with  $e_{\mathfrak{p}} \leq m$  for each  $\mathfrak{p}$ . Since there are at most  $n$  prime ideal factors of  $\mathfrak{p}$ , there are at most  $(m+1)^n$  possibilities for  $\mathfrak{q}$ . For each such  $\mathfrak{q}$  the number of  $\alpha \pmod{p^m}$  with  $\mathfrak{q} \mid (\alpha)$  is  $p^{mn}/N \mathfrak{q} \leq p^{m(n-1)}$ . All together, this yields

$$\varrho(p^m, A) \leq (m+1)^n p^{m(n-1)} \ll m^n p^{m(n-1)} \quad (4.5)$$

whenever  $p^m \mid A$ .

Suppose now that  $p^k \parallel A$  with  $0 \leq k < m$ . Then, for any  $\alpha$  as above,  $p^k \parallel N(\alpha)$  and there is a unique ideal  $\mathfrak{q}$  containing  $\alpha$ , with  $N \mathfrak{q} = p^k$ . Note that  $\mathfrak{q}$  contains  $p^k$ . It follows that

$$\varrho(p^m, A) = \sum_{\mathfrak{q}} \#\{\alpha \pmod{p^m} : \alpha \in \mathfrak{q}, N(\alpha) \equiv A \pmod{p^m}\},$$

where the sum is extended over integral ideals  $\mathfrak{q}$  of norm  $p^k$ . The next goal is to relate, for any of these  $\mathfrak{q}$ , the cardinality above to  $\varrho(p^{m-k}, B)$  for some  $B$  that is coprime to  $p$ . To this end, recall that there exists a prime ideal  $\mathfrak{r}$  in the ideal class  $[\mathfrak{q}]$  which is coprime to  $(p)$ . Suppose that  $\mathfrak{q}(\beta) = \mathfrak{r}(\gamma)$ , so that  $\alpha \in \mathfrak{q}$  if and only if  $\alpha\beta\gamma^{-1} \in \mathfrak{r}$ . We now have

$$\begin{aligned} & \#\{\alpha \pmod{p^m} : \alpha \in \mathfrak{q}, N(\alpha) \equiv A \pmod{p^m}\} \\ &= \#\{\alpha \pmod{p^m} : \alpha\beta\gamma^{-1} \in \mathfrak{r}, N(\alpha\beta\gamma^{-1}) \equiv AN(\beta\gamma^{-1}) \pmod{p^m N(\beta\gamma^{-1})}\} \\ &= \#\{\nu \pmod{p^m \beta\gamma^{-1}} : \nu \in \mathfrak{r}, N(\nu) \equiv B \pmod{p^{m-k} N \mathfrak{r}}\}, \end{aligned}$$

where  $B = Ap^{-k} N \mathfrak{r}$ , and where we note that  $p^m \beta\gamma^{-1} \in \mathfrak{r}$  since  $p^m \in \mathfrak{q}$ . Since  $N \mathfrak{r} \mid N(v)$  for any  $v \in \mathfrak{r}$ , we can replace the final congruence condition above by  $N(\nu) \equiv B \pmod{p^{m-k}}$ . Now choose a  $\mathbb{Z}$ -basis  $\{\xi_1, \dots, \xi_n\}$  for  $\mathfrak{r}$  and recall the definition (4.2) of the associated norm form  $\mathbf{N}(\mathbf{x}; \mathfrak{r})$ . Then the above counting function is equal to the number of integer vectors  $\mathbf{x} \in \mathbb{Z}^n$  producing distinct  $\nu \pmod{p^m \beta\gamma^{-1}}$  for which  $\mathbf{N}(\mathbf{x}; \mathfrak{r}) \equiv B \pmod{p^{m-k}}$ . Our task therefore falls to counting solutions of  $\mathbf{N}(\mathbf{x}; \mathfrak{r}) \equiv B \pmod{p^{m-k}}$  lying in cosets of a certain lattice. To describe this lattice, note that  $\mathfrak{q} \mid (p^k)$  and therefore  $p^{m-k} \mathfrak{r} \mid p^m \mathfrak{r} \mathfrak{q}^{-1}$ . Further, if  $\nu$  and  $\nu'$  agree modulo  $p^{m-k} \mathfrak{r}$  then we have  $\mathbf{x} \equiv \mathbf{x}' \pmod{p^{m-k}}$  and so  $\mathbf{N}(\mathbf{x}; \mathfrak{r})$

and  $\mathbf{N}(\mathbf{x}'; \mathfrak{r})$  coincide modulo  $p^{m-k}$ . It therefore follows that

$$\begin{aligned} & \#\{\nu \pmod{p^m \beta \gamma^{-1}} : \nu \in \mathfrak{r}, N(\nu) \equiv B \pmod{p^{m-k}}\} \\ &= \frac{N(p^m \mathfrak{r} \mathfrak{q}^{-1})}{N(p^{m-k} \mathfrak{r})} \#\{\mathbf{x} \pmod{p^{m-k}} : \mathbf{N}(\mathbf{x}; \mathfrak{r}) \equiv B \pmod{p^{m-k}}\} \\ &= p^{k(n-1)} \varrho(p^{m-k}, B), \end{aligned}$$

by Lemma 4.1, since  $p \nmid N\mathfrak{r}$ . Observing that the number of ideals  $\mathfrak{q}$  of norm  $p^k$  is just  $r_K(p^k)$ , we have therefore shown that for any prime  $p$ , there exists  $B = Ap^{-k} N\mathfrak{r} \in \mathbb{Z}$  such that  $p \nmid B$  and

$$\varrho(p^m, A) = r_K(p^k) p^{k(n-1)} \varrho(p^{m-k}, B), \quad (4.6)$$

whenever  $0 \leq k < m$ .

It remains to analyse  $\varrho(p^m, A)$  when  $p \nmid A$ . Consider the group homomorphism

$$\theta : (\mathfrak{o}/(p^m))^* \rightarrow (\mathbb{Z}/(p^m))^*$$

that is induced by the norm. Since  $\varrho(p^m, A) = \#\theta^{-1}(A) \leq \#\text{Ker } \theta$ , we proceed by bounding its kernel. Suppose first that  $p > 2$  and let  $g$  be a primitive root for  $p^m$ . Note that the image of  $\theta$  contains the subgroup  $H$  generated by  $n$ th powers of elements of  $(\mathbb{Z}/(p^m))^*$ . Hence

$$[(\mathbb{Z}/(p^m))^* : \text{Im}(\theta)] = \frac{[(\mathbb{Z}/(p^m))^* : H]}{[\text{Im}(\theta) : H]} \leq [(\mathbb{Z}/(p^m))^* : H] = \frac{\varphi(p^m)}{\#H}.$$

But  $\#H = \varphi(p^m)/\gcd(n, \varphi(p^m))$ , which readily implies that  $\text{Im}(\theta)$  has index at most  $n$  in  $(\mathbb{Z}/(p^m))^*$ . When  $p = 2$  we argue similarly, using the fact that elements of  $(\mathbb{Z}/(2^m))^*$  can be expressed uniquely as  $(-1)^u 5^v$  for  $u \in \{1, 2\}$  and  $v \in \{1, \dots, 2^{m-2}\}$ , to deduce that  $[(\mathbb{Z}/(p^m))^* : \text{Im}(\theta)] \leq 2n$ . For any prime power  $p^m$  it therefore follows that we have  $\text{Ker}(\theta) \leq 2n\varphi_K(p^m)/\varphi(p^m)$ , where  $\varphi_K$  is the Euler totient function associated to  $K$ . Hence

$$\varrho(p^m, A) \leq 2n \frac{\varphi_K(p^m)}{\varphi(p^m)} \ll p^{m(n-1)}, \quad (4.7)$$

whenever  $p \nmid A$ . We can be more precise when  $p$  is further assumed to be unramified. Assuming that  $p \nmid AD_K$ , we claim that

$$\frac{\varrho(p^m, A)}{p^{m(n-1)}} = \left(1 - \frac{1}{p}\right)^{-1} \prod_{\mathfrak{p}|p} \left(1 - \frac{1}{N\mathfrak{p}}\right). \quad (4.8)$$

Taking  $G = \mathbf{N}$  and  $\ell = 0$  in Lemma 3.4, we see that it suffices to establish this fact when  $m = 1$ . The strategy is to show that the map  $\theta$  is onto, which immediately implies that  $\varrho(p, A) = \#\text{Ker}(\theta) = \varphi_K(p)/\varphi(p)$ , so that the case  $m = 1$  of (4.8) follows. To show that  $\theta$  is onto we must show that there exists  $\mathbf{x} \in \mathbb{F}_p^n$  such that  $\mathbf{N}(\mathbf{x}) \equiv A \pmod{p}$ . For this we deduce from the Chevalley–Warning theorem (see [32, §I.2.2]) that the number of projective solutions is divisible by  $p$ . Moreover, the number of solutions on the hyperplane at infinity is

$$\#\{\mathbf{x} \in \mathbb{F}_p^n : \mathbf{N}(\mathbf{x}) \equiv 0 \pmod{p}\} = p^n - \#\{\mathbf{x} \in \mathbb{F}_p^n : p \nmid \mathbf{N}(\mathbf{x})\} = p^n - \varphi_K(p).$$

Since  $p \nmid \varphi_K(p)$  for an unramified prime  $p$ , we may conclude that the number of affine solutions to the congruence  $\mathbf{N}(\mathbf{x}) \equiv A \pmod{p}$  is not divisible by  $p$ . This shows that  $\theta$  is onto, as required.

We may now conclude the proof of Lemma 4.2. The second part follows from (4.6) and (4.8). Recalling from (2.10) that  $r_K(p^k) \leq (k+1)^n$ , the first part follows from (4.5), (4.6) and (4.7).  $\square$

## 5. COUNTING POINTS ON SYSTEMS OF NORM FORM EQUATIONS

While the previous two sections described background, notation and technical tools, we now begin with the proof of our main theorem. In the first two parts of this section we state and discuss our main auxiliary result which may be interpreted as an asymptotic formula for the number of integral points of bounded height on an integral model for the variety  $\mathcal{V} \subset \mathbb{A}_{\mathbb{Q}}^{n_1+\dots+n_r+s}$  defined in (1.3). In the final part of this section we deduce Theorem 1.3 from this asymptotic formula.

**5.1. Representation function and asymptotic formula.** After a change of variables we may assume that we are working with an integral model for  $\mathcal{V}$ , defined by the system of equations

$$0 \neq \mathbf{N}_{K_i}(\mathbf{x}_i) = f_i(u_1, \dots, u_s), \quad (1 \leq i \leq r),$$

where each  $K_i$  is a number field of degree  $n_i > 1$ , each  $f_i$  is a linear form defined over  $\mathbb{Z}$ , and the forms  $f_i$  are pairwise non-proportional. We further assume that each  $\mathbf{N}_{K_i}$  is defined using a  $\mathbb{Z}$ -basis  $\{\omega_{i,1}, \dots, \omega_{i,n_i}\}$  for the ring of integers of  $\mathfrak{o}_{K_i}$ , so that it too has integer coefficients.

We will phrase the problem of counting integral points on  $\mathcal{V}$  in terms of representation functions  $R_i : \mathbb{Z} \rightarrow \mathbb{Z}_{\geq 0}$  that, in the simplest instance, count the number of representations  $m = \mathbf{N}_{K_i}(\mathbf{x}_i)$  of each non-zero integer  $m$ , where  $\mathbf{x}_i$  runs through equivalence classes with respect to the action of the free part  $Y_{K_i}^{(+)}$  of  $U_{K_i}^{(+)}$ . Our application to Theorem 1.3 requires us to incorporate some flexibility into the definition of  $R_i$  as to exactly which representations are counted. To describe these restrictions, we use the notation of Section 2. In particular, recall that

$$\mathfrak{D}_{i,+} = \{\mathbf{x} \in \mathbb{R}^{n_i} : x_1\omega_{i,1} + \dots + x_{n_i}\omega_{i,n_i} \in \mathfrak{F}_{i,+}\}$$

is a fundamental domain for the action of  $Y_{K_i}^{(+)}$  on the coordinate space  $\mathbb{R}^{n_i}$ . Furthermore, we recall from (2.4) that

$$\mathfrak{D}_{i,+}^{\epsilon}(T) = \{\mathbf{x} \in \mathfrak{D}_{i,+} : 0 < \epsilon \mathbf{N}_{K_i}(\mathbf{x}) \leq T\},$$

for  $\epsilon \in \{\pm\}$  and  $T \geq 1$ .

**Definition 5.1** (Representation function). Let  $i \in \{1, \dots, r\}$  and let  $\mathfrak{X}_i \subset \mathfrak{D}_{i,+}$  be a cone such that each of the bounded sets  $\mathfrak{X}_i \cap \mathfrak{D}_{i,+}^{\epsilon}(1)$  has an  $(n_i - 1)$ -Lipschitz parametrisable

boundary, unless it is empty. Let  $M \in \mathbb{Z}_{>0}$  and let  $\mathbf{b}_i \in (\mathbb{Z}/M\mathbb{Z})^{n_i}$  for  $1 \leq i \leq r$ . For any  $m \in \mathbb{Z}$  we define

$$R_i(m; \mathfrak{X}_i, \mathbf{b}_i, M) = \mathbf{1}_{m \neq 0} \cdot \# \left\{ \mathbf{x} \in \mathbb{Z}^{n_i} \cap \mathfrak{X}_i : \begin{array}{l} \mathbf{N}_{K_i}(\mathbf{x}) = m \\ \mathbf{x} \equiv \mathbf{b}_i \pmod{M} \end{array} \right\}.$$

We shall abbreviate  $R_i(m) = R_i(m; \mathfrak{X}_i, \mathbf{b}_i, M)$ , once  $\mathfrak{X}_i, \mathbf{b}_i$  and  $M$  are fixed.

Next, let  $\mathfrak{K} \subset \mathbb{R}^s$  be any convex bounded set. Our interest lies in the counting function

$$N(T) = \sum_{\substack{\mathbf{u} \in \mathbb{Z}^s \cap T\mathfrak{K} \\ \mathbf{u} \equiv \mathbf{a} \pmod{M}}} \prod_{i=1}^r R_i(f_i(\mathbf{u})), \quad (5.1)$$

for given  $\mathbf{a} \in (\mathbb{Z}/M\mathbb{Z})^s$ . By unravelling the definition of  $R_i$ , this is seen to express the number of suitably constrained points in  $\mathcal{V}(\mathbb{Z})$ . For technical reasons, we restrict attention to  $\mathbf{a} \in (\mathbb{Z}/M\mathbb{Z})^s$  such that  $p^{v_p(M)} \nmid f_i(\mathbf{a})$  for any  $p \mid M$  and any  $1 \leq i \leq r$ .

From now on we will view  $M, K_1, \dots, K_r$ , together with  $\mathbf{a}, \mathbf{b}_i$  and the coefficients of  $\mathcal{V}$  as being fixed once and for all. Any implied constants in our work will therefore be allowed to depend on these quantities in any way. Moreover, the regions  $\mathfrak{X}_1, \dots, \mathfrak{X}_r$  are also to be considered fixed, with any implied constant being allowed to depend on the Lipschitz constants of the maps parametrising the boundaries.

Before revealing our asymptotic formula for  $N(T)$  we require a bit more notation. For given  $q \in \mathbb{Z}_{>0}$  and  $A \in \mathbb{Z}$ , with  $M \mid q$ , we let

$$\varrho_i(q, A; M) = \# \left\{ \mathbf{x} \in (\mathbb{Z}/q\mathbb{Z})^{n_i} : \begin{array}{l} \mathbf{N}_{K_i}(\mathbf{x}) \equiv A \pmod{q} \\ \mathbf{x} \equiv \mathbf{b}_i \pmod{M} \end{array} \right\},$$

for  $1 \leq i \leq r$ , as in (4.4). Moreover, for  $\epsilon \in \{\pm\}$ , we define

$$\mathbb{R}_\epsilon = \{x \in \mathbb{R} : \epsilon x > 0\}$$

and

$$\kappa_i^\epsilon(\mathfrak{X}_i) = \text{vol}(\mathfrak{D}_{i,+}^\epsilon(1) \cap \mathfrak{X}_i). \quad (5.2)$$

Finally, we denote by  $\mathbf{f} : \mathbb{R}^s \rightarrow \mathbb{R}^r$  the linear map defined by the system  $\mathbf{f} = (f_1, \dots, f_r)$  of linear forms. Bearing this notation in mind we have the following result.

**Theorem 5.2.** *Let  $f_1, \dots, f_r \in \mathbb{Z}[u_1, \dots, u_s]$  be pairwise non-proportional linear forms and assume that  $|f_i(\mathfrak{K})| \leq 1$ , for  $1 \leq i \leq r$ . Suppose that  $M, \mathbf{b}_i$  and  $\mathbf{a}$  are as above; in particular,  $p^{v_p(M)} \nmid f_i(\mathbf{a})$  for any  $p \mid M$  and any  $1 \leq i \leq r$ . Then we have*

$$N(T) = \beta_\infty \prod_p \beta_p \cdot T^s + o(T^s), \quad (T \rightarrow \infty),$$

where

$$\beta_\infty = \sum_{\epsilon \in \{\pm\}^r} \text{vol}(\mathfrak{K} \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r})) \prod_{i=1}^r \kappa_i^{\epsilon_i}(\mathfrak{X}_i)$$

and

$$\beta_p = \lim_{m \rightarrow \infty} \frac{1}{p^{ms}} \sum_{\substack{\mathbf{u} \in (\mathbb{Z}/p^m \mathbb{Z})^s \\ \mathbf{u} \equiv \mathbf{a} \pmod{p^{v_p(M)}}}} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}); p^{v_p(M)})}{p^{m(n_i-1)}},$$

for each prime  $p$ . Furthermore, the product  $\prod_p \beta_p$  is absolutely convergent.

We will show how Theorem 5.2 implies Theorem 1.3 in Section 5.3. The proof of Theorem 5.2 takes up most of the remainder of this paper. The first part is established in the course of Sections 6–10, while the final part is dealt with in Section 5.2 below.

**Remark 5.3.** Our proof uses the machinery developed in Green and Tao [18]. As such, it in fact covers the case where in the statement of Theorem 5.2 each linear form  $f_i$  is replaced by a linear polynomial  $f_i + a_i$ , for an integer  $a_i = O(T)$ .

**Remark 5.4.** In the special case where  $\mathfrak{X}_i = \mathfrak{D}_{i,+}$  for  $1 \leq i \leq r$  and  $M = 1$  in  $N(T)$ , it is straightforward to adapt the calculation in [25, §6] to find a precise value for  $\kappa_i^\epsilon(\mathfrak{D}_{i,+}) = \text{vol}(\mathfrak{D}_{i,+}^\epsilon(1))$ . Let us drop the index  $i$  and work with a typical field  $K$  of degree  $n$ . Let  $\delta_1, \dots, \delta_{r_1+r_2-1}$  be generators for  $Y_K^{(+)}$ . We define a modified regulator  $R_K^{(+)}$  to be the absolute value of the determinant of the  $(r_1 + r_2) \times (r_1 + r_2)$  matrix, whose rows are given by  $\psi(\delta_1), \dots, \psi(\delta_{r_1+r_2-1}), \mathbf{u}_{r_1+r_2} \in \mathbb{R}^{r_1+r_2}$ , in the notation of Section 2.2. Then one finds that

$$\kappa_i^\epsilon(\mathfrak{D}_{i,+}) = \begin{cases} 0, & \text{if } \epsilon = - \text{ and } r_1 = 0, \\ 2^{r_1-1} (2\pi)^{r_2} R_K^{(+)} / \sqrt{|D_K|}, & \text{otherwise.} \end{cases}$$

Observing that  $\psi(\eta_1^2) = 2\psi(\eta_1)$ , furthermore, an inspection of the explicit choice of generators for  $Y_K^{(+)}$  given in Section 2.2 shows that  $R_K^{(+)} = [Y_K : Y_K^{(+)}] R_K$ . Theorem 5.2 recovers [28, Thm. 1.1] when  $K_1, \dots, K_r$  are all taken to be quadratic.

**5.2. Convergence of the product of local densities.** In this section we prove the absolute convergence of the product  $\prod_p \beta_p$  from Theorem 5.2, by establishing an asymptotic estimate for the local density  $\beta_p$ , valid whenever  $p$  is large compared to

$$L = \max_{1 \leq i \leq r} \{\|f_i\|, s, r, |D_{K_i}|\} \quad (5.3)$$

and  $p \nmid M$ . Here  $\|f_i\|$  denotes the maximum modulus of the coefficients of  $f_i$ .

**Proposition 5.5.** *We have  $\beta_p = 1 + O_L(p^{-2})$  whenever  $p \nmid M$  and  $\beta_p = O_L(1)$  when  $p|M$ . In particular, there exists  $L' = O_L(1)$ , which is independent of  $M$ , such that  $\beta_p > 0$  whenever  $p > L'$  and  $p \nmid M$ .*

This proposition immediately implies the convergence of the product  $\prod_p \beta_p$ . The proof of Proposition 5.5 splits into two cases according to whether  $p$  is large or small compared to  $L$ , and follows that of [27, Lemma 8.3]. The main ingredients are the information that Lemma 4.2 provides about  $\varrho(q, A)$ , and the properties of local divisor densities, which we discuss next.

Let

$$\mathcal{U}_m = \{\mathbf{u} \in (\mathbb{Z}/p^m\mathbb{Z})^s : \mathbf{u} \equiv \mathbf{a} \pmod{p^{v_p(M)}}\}, \quad (5.4)$$

for any  $m \in \mathbb{Z}_{>0}$ . For given  $\mathbf{c} \in \mathbb{Z}_{\geq 0}^r$  and a given system  $\mathbf{f} = (f_1, \dots, f_r)$  as above, we define the *local divisor density* (cf. [18, p.1831] and [27, Def. 8.4]) to be

$$\alpha_{\mathbf{f}}(p^{c_1}, \dots, p^{c_r}) = \frac{1}{p^{ms}} \sum_{\mathbf{u} \in \mathcal{U}_m} \prod_{i=1}^r \mathbf{1}_{p^{c_i} | f_i(\mathbf{u})}, \quad (5.5)$$

where  $m = \max\{c_1, \dots, c_r\}$ . Let  $n(\mathbf{c})$  denote the number of non-zero components of  $\mathbf{c}$ . Then

$$\alpha_{\mathbf{f}}(p^{c_1}, \dots, p^{c_r}) \begin{cases} = 1, & \text{if } n(\mathbf{c}) = 0, \\ = p^{-\max_i \{c_i\}}, & \text{if } p \gg_L 1, p \nmid M \text{ and } n(\mathbf{c}) = 1, \\ \leq p^{-\max_{i \neq j} \{c_i + c_j\}}, & \text{if } p \gg_L 1, p \nmid M \text{ and } n(\mathbf{c}) > 1, \\ \ll_L p^{-\max_i \{c_i\}}, & \text{otherwise.} \end{cases} \quad (5.6)$$

It is important to note here that even when  $p \mid M$  and  $n(\mathbf{c}) \geq 1$  the implied constant in the final estimate does not depend on  $M$ . Moreover, here (and elsewhere) we take  $p \gg_L 1$  to mean that  $p$  is sufficiently large in terms of  $L$ . An easy way to bound sums over divisor densities uses the observation that there are at most  $rJ^{r-1}$  choices of  $\mathbf{k} \in \mathbb{Z}_{\geq 0}^r$  such that  $\max_i k_i = J$  and therefore

$$\sum_{J \geq J_0} \sum_{\substack{\mathbf{k} \in \mathbb{Z}_{\geq 0}^r \\ \max\{k_1, \dots, k_r\} = J}} \frac{J^T}{p^J} \leq r \sum_{J \geq J_0} \frac{J^{T+r-1}}{p^J} \ll_{T,r,J_0} \frac{1}{p^{J_0}} \sum_{J \geq 0} \frac{J^{T+r-1}}{2^J} \ll_{T,r,J_0} \frac{1}{p^{J_0}} \quad (5.7)$$

for any  $T, J_0 > 0$ .

*Proof of Proposition 5.5.* We may write  $\beta_p = \lim_{m \rightarrow \infty} \beta_p(m)$ , with

$$\beta_p(m) = \frac{1}{p^{ms}} \sum_{\mathbf{k} \in \mathbb{Z}_{\geq 0}^r} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ v_p(f_i(\mathbf{u})) = k_i}} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}); p^\mu)}{p^{m(n_i-1)}} \quad (5.8)$$

and  $\mu = v_p(M)$ . We begin by analysing  $\beta_p$  when  $p$  is small. In fact we will show that  $\beta_p = O_L(1)$ , for any prime  $p$ , which suffices for Proposition 5.5.

Since

$$\varrho_i(p^m, f_i(\mathbf{u}); p^\mu) \leq \varrho_i(p^m, f_i(\mathbf{u})),$$

an application of the first part of Lemma 4.2 in (5.8) shows that

$$\begin{aligned} \beta_p(m) &\leq \frac{1}{p^{ms}} \sum_{J \geq 0} \sum_{\substack{\mathbf{k} \in \mathbb{Z}_{\geq 0}^r \\ \max\{k_1, \dots, k_r\} = J}} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ v_p(f_i(\mathbf{u})) = k_i}} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}))}{p^{m(n_i-1)}} \\ &\ll \frac{1}{p^{ms}} \sum_{J \geq 0} \sum_{\substack{\mathbf{k} \in \mathbb{Z}_{\geq 0}^r \\ \max\{k_1, \dots, k_r\} = J}} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ p^{k_i} | f_i(\mathbf{u})}} \min\{m, J+1\}^{n_1 + \dots + n_r}. \end{aligned}$$

Next we invoke (5.6) and (5.7) to obtain

$$\beta_p(m) \ll_L \sum_{J \geq 0} \sum_{\substack{\mathbf{k} \in \mathbb{Z}_{\geq 0}^r \\ \max\{k_1, \dots, k_r\} = J}} \frac{\min\{m, J+1\}^{n_1 + \dots + n_r}}{p^J} \ll_L 1.$$

Taking the limit  $m \rightarrow \infty$ , this shows that  $\beta_p = O_L(1)$ , as required for Proposition 5.5.

We proceed to analyse  $\beta_p$  when  $p \gg_L 1$  and  $p \nmid M$ . In particular, we have  $\mu = 0$  and  $\mathcal{U}_m = (\mathbb{Z}/p^m\mathbb{Z})^s$ . Let  $\mathcal{K}_1 = (\mathbb{Z} \cap [0, m))^r$  and let  $\mathcal{K}_2 = \mathbb{Z}_{\geq 0}^r \setminus [0, m)^r$ . Accordingly, we write  $\beta_p(m) = \beta_p^{(1)}(m) + \beta_p^{(2)}(m)$ , where  $\beta_p^{(i)}(m)$  is the contribution from  $\mathbf{k} \in \mathcal{K}_i$ .

Since  $p \gg_L 1$ , it follows from (5.3) that  $p \nmid D_{K_i}$  for each  $1 \leq i \leq r$ . Thus the second part of Lemma 4.2 implies that

$$\begin{aligned} \beta_p^{(1)}(m) &= c_p(K_1) \dots c_p(K_r) \frac{1}{p^{ms}} \sum_{\mathbf{k} \in \mathcal{K}_1} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ v_p(f_i(\mathbf{u})) = k_i}} \prod_{i=1}^r r_{K_i}(p^{k_i}) \\ &= c_p(K_1) \dots c_p(K_r) \frac{1}{p^{ms}} \sum_{\mathbf{k} \in \mathcal{K}_1} \left( \prod_{i=1}^r r_{K_i}(p^{k_i}) \right) \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ v_p(f_i(\mathbf{u})) = k_i}} 1, \end{aligned}$$

where

$$c_p(K_i) = \left(1 - \frac{1}{p}\right)^{-1} \prod_{\mathfrak{p} \mid p, \mathfrak{p} \subset \mathfrak{o}_{K_i}} \left(1 - \frac{1}{N\mathfrak{p}}\right).$$

For given  $\mathbf{k} \in \mathcal{K}_1$ , we have

$$\begin{aligned} \frac{1}{p^{ms}} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ v_p(f_i(\mathbf{u})) = k_i}} 1 &= \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^r} (-1)^{\varepsilon_1 + \dots + \varepsilon_r} \alpha_{\mathbf{f}}(p^{k_1 + \varepsilon_1}, \dots, p^{k_r + \varepsilon_r}) \\ &= \begin{cases} 1 - rp^{-1} + O_r(p^{-2}), & \text{if } \mathbf{k} = \mathbf{0}, \\ p^{-1} + O_r(p^{-2}), & \text{if } \mathbf{k} \in \{(1, 0, \dots, 0), \dots, (0, \dots, 0, 1)\}, \\ O_r(p^{-\max\{k_1, \dots, k_r\}}), & \text{if } n(\mathbf{k}) = 1, \max\{k_1, \dots, k_r\} > 1, \\ O_r(p^{-1 - \max\{k_1, \dots, k_r\}}), & \text{otherwise,} \end{cases} \end{aligned}$$

by (5.6).

Since  $r_{K_i}(p^k) = O((k+1)^{n_i})$ , by (2.10), we deduce from (5.7) that

$$\sum_{J \geq 2} \sum_{\substack{\mathbf{k} \in \mathcal{K}_1 \\ \max\{k_1, \dots, k_r\} = J \\ n(\mathbf{k}) = 1}} p^{-J} \prod_{i=1}^r r_{K_i}(p^{k_i}) + \sum_{J \geq 1} \sum_{\substack{\mathbf{k} \in \mathcal{K}_1 \\ \max\{k_1, \dots, k_r\} = J \\ n(\mathbf{k}) > 1}} p^{-J-1} \prod_{i=1}^r r_{K_i}(p^{k_i}) \ll_r \frac{1}{p^2}.$$

Hence (2.11) implies that

$$\begin{aligned}\beta_p^{(1)}(m) &= c_p(K_1) \dots c_p(K_r) \left( 1 + \sum_{i=1}^r \frac{r_{K_i}(p) - 1}{p} + O_r \left( \frac{1}{p^2} \right) \right) \\ &= 1 + O_r \left( \frac{1}{p^2} \right).\end{aligned}$$

Putting everything together, we conclude that

$$|\beta_p(m) - 1| \ll_r \frac{1}{p^2} + \beta_p^{(2)}(m).$$

The first part of Lemma 4.2 can be used to show that  $\beta_p^{(2)}(m)$  is at most

$$\begin{aligned}\frac{1}{p^{ms}} \sum_{\mathbf{k} \in \mathcal{K}_2} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ p^{k_i} | f_i(\mathbf{u})}} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}))}{p^{m(n_i-1)}} &\ll \frac{m^{n_1+\dots+n_r}}{p^{ms}} \sum_{\mathbf{k} \in \mathcal{K}_2} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m \\ p^{k_i} | f_i(\mathbf{u})}} 1 \\ &= m^{n_1+\dots+n_r} \sum_{\mathbf{k} \in \mathcal{K}_2} \alpha_{\mathbf{f}}(p^{k_1}, \dots, p^{k_r}) \\ &\ll \frac{m^{n_1+\dots+n_r}}{p^{1+m}},\end{aligned}$$

by (5.6). Substituting this into our expression for  $\beta_p$  and taking the limit  $m \rightarrow \infty$ , this completes the proof of Proposition 5.5 when  $p \gg_L 1$ .  $\square$

**5.3. Deduction of Theorem 1.3.** We proceed to show how Theorem 1.3 follows from Theorem 5.2. Our task is to establish the Hasse principle and weak approximation for the smooth variety  $\mathcal{V} \subset \mathbb{A}_{\mathbb{Q}}^{n_1+\dots+n_r+s}$ , which after the reductions from the start of Section 5.1 is given by

$$0 \neq \mathbf{N}_{K_i}(\mathbf{x}_i) = f_i(u_1, \dots, u_s), \quad (1 \leq i \leq r),$$

for pairwise non-proportional linear forms  $f_1, \dots, f_r$  defined over  $\mathbb{Z}$ .

Suppose that we are given a point  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{Q})$ . Then each point in the orbit  $\{(\mathbf{u}, \eta_i \cdot \mathbf{x}_i) : \eta_i \in U_{K_i}^{(+)}\}$  also belongs to  $\mathcal{V}(\mathbb{Q})$ . We will therefore content ourselves with looking for points  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{Q})$  such that each  $\mathbf{x}_i$  lies in the fundamental domain  $\mathfrak{D}_{i,+}$  which we constructed in Lemma 2.1 and (2.3) for the free part of  $U_{K_i}^{(+)}$ . We will call such points *primary*.

Let  $\Omega$  denote the set of places of  $\mathbb{Q}$ . We assume we are given points  $(\mathbf{u}^{(\nu)}, \mathbf{x}_i^{(\nu)}) \in \mathcal{V}(\mathbb{Q}_{\nu})$  for every  $\nu \in \Omega$ . By possibly replacing the adelic point  $(\mathbf{x}_i^{(\nu)})_{\nu \in \Omega}$  by  $(\eta_i \cdot \mathbf{x}_i^{(\nu)})_{\nu \in \Omega}$  for an appropriate  $\eta_i \in U_{K_i}^{(+)}$ , we may assume that  $\mathbf{x}_i^{(\infty)}$  belongs to  $\mathfrak{D}_{i,+}$  for each  $1 \leq i \leq r$ . Let  $S$  be any finite set of places, including the archimedean place as well as all non-archimedean places corresponding to primes  $p < L'$ , where  $L' = O(1)$  was determined in Proposition 5.5.



Let  $\varepsilon > 0$ . Then, in order to prove Theorem 1.3, it suffices to show that there is a primary point  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{Q})$  such that

$$|\mathbf{u} - \mathbf{u}^{(\nu)}|_\nu < \varepsilon, \quad |\mathbf{x}_i - \mathbf{x}_i^{(\nu)}|_\nu < \varepsilon, \quad (1 \leq i \leq r), \quad (5.9)$$

for every  $\nu \in S$ . Here,  $|\cdot|_\nu$  denotes the  $\nu$ -adic norm extended to vectors in the obvious way, and we follow the convention that  $|\cdot|_\infty = |\cdot|$ .

On rescaling appropriately we may assume that the points  $(\mathbf{u}^{(\nu)}, \mathbf{x}_i^{(\nu)})$  that we are given belong to  $\mathbb{Z}_\nu^{n_1 + \dots + n_r + s}$  for every finite  $\nu \in S$ . By the Chinese remainder theorem we can then produce an integer vector  $(\mathbf{u}^{(M)}, \mathbf{x}_i^{(M)})$  such that

$$|\mathbf{u}^{(M)} - \mathbf{u}^{(\nu)}|_\nu < \varepsilon, \quad |\mathbf{x}_i^{(M)} - \mathbf{x}_i^{(\nu)}|_\nu < \varepsilon, \quad (1 \leq i \leq r), \quad (5.10)$$

for all finite  $\nu \in S$ . We now seek integral points  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{Z})$  satisfying the following local conditions. For the finite places we impose

$$\mathbf{u} \equiv \mathbf{u}^{(M)} \pmod{M}, \quad \mathbf{x}_i \equiv \mathbf{x}_i^{(M)} \pmod{M}, \quad (1 \leq i \leq r), \quad (5.11)$$

for an appropriate modulus  $M \in \mathbb{Z}_{>0}$  with the property that  $p \nmid M$  whenever  $p \notin S$ . In view of (5.10) these conditions imply (5.9). To guarantee that

$$p^{v_p(M)} \nmid f_i(\mathbf{u}^{(M)}) \text{ for all } p \mid M, \quad (1 \leq i \leq r), \quad (5.12)$$

it suffices to choose  $\varepsilon$  sufficiently small, since  $f_i(\mathbf{u}^{(\nu)}) \neq 0$  in  $\mathbb{Q}_\nu$ .

For the infinite place we impose that

$$|\mathbf{u} - B\mathbf{u}^{(\infty)}| < \varepsilon B, \quad |\mathbf{x}_i - B^{1/n_i}\mathbf{x}_i^{(\infty)}| < \varepsilon B^{1/n_i}, \quad (1 \leq i \leq r), \quad (5.13)$$

with  $B = P^{n_1 \dots n_r}$  and  $P \in \mathbb{Z}_{>0}$  tending to infinity such that  $P \equiv 1 \pmod{M}$ . Thus any point  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{Z})$  satisfying (5.11) and (5.13) gives rise to  $(B^{-1}\mathbf{u}, B^{-1/n_i}\mathbf{x}_i) \in \mathcal{V}(\mathbb{Q})$  satisfying the original condition (5.9). We aim to detect the existence of integral points satisfying (5.11) and (5.13) using Theorem 5.2. For this reason, we now proceed to replace (5.13) by a condition that is more suitable for an application of the theorem.

Let  $1 \leq i \leq r$  and let  $\varepsilon' > 0$ . We begin by defining a cone that is symmetric about  $\mathbf{x}_i^{(\infty)}$ , via

$$\mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon') = \left\{ \mathbf{x} \in \mathbb{R}^{n_i} \cap \mathfrak{D}_{i,+} : \left| \frac{\mathbf{x}}{|\mathbf{N}_{K_i}(\mathbf{x})|^{1/n_i}} - \frac{\mathbf{x}_i^{(\infty)}}{|\mathbf{N}_{K_i}(\mathbf{x}_i^{(\infty)})|^{1/n_i}} \right| < \varepsilon' \right\}.$$

Note that  $\mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon') \neq \emptyset$ , since  $\mathbf{x}_i^{(\infty)} \in \mathfrak{D}_{i,+}$  by our work above. It follows from Sections 2.2 and 3.1, that for  $\epsilon \in \{\pm\}$  each  $\mathfrak{D}_{i,+}^\epsilon(1) \cap \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon')$  is either empty or such that Lemma 3.1 applies. Indeed, these sets arise as the intersection of a bounded convex set with a set that has an  $(n_i - 1)$ -Lipschitz parametrisable boundary. Moreover, we clearly have

$$\text{vol}(\mathfrak{D}_{i,+}^\epsilon(1) \cap \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon')) > 0 \quad (5.14)$$

when  $\epsilon = \text{sign}(\mathbf{N}_{K_i}(\mathbf{x}_i^{(\infty)}))$ . The second condition in (5.13) now holds whenever

$$\mathbf{x}_i \in \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon') \cap \{\mathbf{x} \in \mathbb{R}^{n_i} : |\mathbf{N}_{K_i}(\mathbf{x}) - B\mathbf{N}_{K_i}(\mathbf{x}_i^{(\infty)})| < \varepsilon' B\} \quad (5.15)$$

for sufficiently small  $\varepsilon'$  in terms of  $\max_{1 \leq i \leq r} |\mathbf{N}_{K_i}(\mathbf{x}_i^{(\infty)})|$ . We fix such a choice of  $\varepsilon' < \varepsilon$ .

In view of the first part of (5.13), we define the convex bounded region

$$\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'') = \{\mathbf{u} \in \mathbb{R}^s : |\mathbf{u} - \mathbf{u}^{(\infty)}| < \varepsilon''\},$$

for  $\varepsilon'' > 0$ . Observe that for sufficiently small  $\varepsilon'' < \varepsilon$ , the condition  $\mathbf{u} \in B\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'')$  implies both the first part of (5.13) and, furthermore,

$$|f_i(\mathbf{u}) - Bf_i(\mathbf{u}^{(\infty)})| < \varepsilon' B.$$

In conclusion, any point  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{R})$  with

$$\mathbf{u} \in B\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'') \quad \text{and} \quad \mathbf{x}_i \in \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon') \quad (5.16)$$

satisfies (5.15) and therefore also (5.13).

With these choices of  $\varepsilon', \varepsilon''$ , we fix the representation functions

$$R_i(m) = R_i(m; \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon'), \mathbf{x}_i^{(M)}, M)$$

from Definition 5.1, for  $1 \leq i \leq r$ . It is clear that

$$|f_i(\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon''))| \subset [-H, H]$$

for  $H = s(1 + \varepsilon)|\mathbf{u}^{(\infty)}| \cdot \max_i \|f_i\|$ . In particular,  $H \asymp 1$ . Thus we observe that, on the one hand,

$$N(HB) = \sum_{\substack{\mathbf{u} \in \mathbb{Z}^s \cap HB(H^{-1}\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'')) \\ \mathbf{u} \equiv \mathbf{u}^{(M)} \pmod{M}}} \prod_{i=1}^r R_i(f_i(\mathbf{u}))$$

counts exactly the primary points  $(\mathbf{u}, \mathbf{x}_i) \in \mathcal{V}(\mathbb{Z})$  which satisfy (5.11) and (5.16). On the other hand,  $N(HB)$  takes the shape of the counting function (5.1) from Theorem 5.2, with

$$\mathfrak{K} = H^{-1}\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon''), \quad \mathfrak{X}_i = \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon'), \quad \mathbf{a} = \mathbf{u}^{(M)}, \quad \mathbf{b}_i = \mathbf{x}_i^{(M)}.$$

Moreover, all the conditions of Theorem 5.2 are satisfied. We conclude, for  $T = HB$ , that

$$N(HB) = (HB)^s \beta_\infty \prod_p \beta_p + o(B^s),$$

where the error term may depend on  $\varepsilon', \varepsilon'', M, K_1, \dots, K_r$ , as well as on  $\mathbf{u}^{(M)}, \mathbf{u}^{(\infty)}, \mathbf{x}_i^{(M)}, \mathbf{x}_i^{(\infty)}$  and on the coefficients of  $\mathcal{V}$ . All that remains now, in order to deduce Theorem 1.3, is to show that

$$\beta_\infty \prod_p \beta_p \gg 1.$$

Beginning with  $\beta_\infty$  we recall that

$$\kappa_i^\epsilon = \kappa_i^\epsilon(\mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon')) = \text{vol} \left( \mathfrak{D}_{i,+}^\epsilon(1) \cap \mathfrak{B}_i(\mathbf{x}_i^{(\infty)}; \varepsilon') \right).$$

Thus, (5.14) yields  $\kappa_i^{\epsilon_i} > 0$  when  $\epsilon_i = \text{sign}(f_i(\mathbf{u}^{(\infty)}))$  for each  $1 \leq i \leq r$ . Next we check that for these choices of  $\epsilon_1, \dots, \epsilon_r$  we also have

$$\text{vol}(H^{-1}\mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'') \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r})) > 0.$$

Since  $f_1, \dots, f_r$  are linear homogeneous polynomials, the region  $\mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r})$  is a cone. Thus

$$\begin{aligned} \text{vol} \left( H^{-1} \mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'') \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r}) \right) \\ = H^{-s} \text{vol} \left( \mathfrak{K}(\mathbf{u}^{(\infty)}; \varepsilon'') \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r}) \right), \end{aligned}$$

which is positive, since  $\mathbf{u}^{(\infty)}$  is an element of the open set  $\mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r})$ .

Turning to the local factors at the non-archimedean places  $p$ , we recall that

$$\beta_p = \lim_{m \rightarrow \infty} \frac{1}{p^{ms}} \sum_{\mathbf{u} \in \mathcal{U}_m} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}); p^{v_p(M)})}{p^{m(n_i-1)}},$$

where  $\mathcal{U}_m$  is given by (5.4), with  $\mathbf{a} = \mathbf{u}^{(M)}$ . By construction, we have  $p \nmid M$  whenever  $p \notin S$  and it follows from Proposition 5.5 that

$$\prod_{p \notin S} \beta_p \gg 1,$$

and that  $\beta_p = O(1)$  for  $p \in S$ . This leaves us to show that  $\beta_p > 0$  for every  $p \in S$  in order to complete the proof of Theorem 1.3. We will deduce this with the help of Lemma 3.4 from the existence of local solutions at these primes. While we are primarily interested in  $p \in S$ , the following argument works for any prime  $p$ .

With Lemma 3.4 in mind, we fix a prime  $p$  and let

$$m' = 2 \left( 1 + v_p(M) + \sum_{i=1}^r v_p(f_i(\mathbf{u}^{(M)})) + \sum_{i=1}^r v_p(n_i) \right).$$

Recall that we are given  $(\mathbf{u}^{(p)}, \mathbf{x}_i^{(p)}) \in \mathcal{V}(\mathbb{Q}_p)$  such that (5.10) holds for  $\nu = p$ . By solving this approximation problem for a smaller value of  $\varepsilon$ , we can find  $(\mathbf{u}', \mathbf{x}'_i) \in \mathbb{Z}^{n_1 + \dots + n_r + s}$  such that

$$(\mathbf{u}', \mathbf{x}'_i) \equiv (\mathbf{u}^{(M)}, \mathbf{x}_i^{(M)}) \pmod{p^{v_p(M)}} \quad \text{and} \quad f_i(\mathbf{u}') \equiv \mathbf{N}_{K_i}(\mathbf{x}'_i) \pmod{p^{m'}}.$$

Thus,

$$\prod_{i=1}^r \varrho_i(p^{m'}, f_i(\mathbf{u}'); p^{v_p(M)}) \geq 1,$$

where each  $\varrho_i$  is defined with respect to  $\mathbf{x}_i^{(M)}$ . The technical condition (5.12) ensures that such approximations  $\mathbf{u}'$  satisfy  $v_p(f_i(\mathbf{u}')) = v_p(f_i(\mathbf{u}^{(M)}))$  and  $f_i(\mathbf{u}') \neq 0$ .

The definition of  $m'$  ensures that the conditions of Lemma 3.4 are satisfied when  $m \geq m'$ ,  $\ell = v_p(M)$ ,  $A = f_i(\mathbf{u}')$ ,  $\mathbf{a} = \mathbf{x}_i^{(M)}$  and  $G = \mathbf{N}_{K_i}$ , for any  $1 \leq i \leq r$ . Hence we obtain

$$\begin{aligned} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}); p^{v_p(M)})}{p^{m(n_i-1)}} &= \prod_{i=1}^r \frac{\varrho_i(p^{m'}, f_i(\mathbf{u}'); p^{v_p(M)})}{p^{m'(n_i-1)}} \\ &\geq \frac{1}{p^{m'(n_1 + \dots + n_r - r)}}, \end{aligned}$$

whenever  $\mathbf{u} \in \mathcal{U}'_m = \{\mathbf{u} \in (\mathbb{Z}/p^m\mathbb{Z})^s : \mathbf{u} \equiv \mathbf{u}' \pmod{p^{m'}}\}$ . The set  $\mathcal{U}'_m$  has  $p^{(m-m')s}$  elements and is clearly a subset of  $\mathcal{U}_m$ . Therefore

$$\begin{aligned} \beta_p(m) &\geq \frac{1}{p^{ms}} \sum_{\mathbf{u} \in \mathcal{U}'_m} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}); p^{v_p(M)})}{p^{m(n_i-1)}} \\ &\geq \frac{1}{p^{m'(n_1+\dots+n_r+s-r)}}, \end{aligned}$$

for every  $m \geq m'$ , which provides the desired inequality  $\beta_p > 0$  for  $p \in S$ .

## 6. $W$ -TRICK AND NON-CORRELATION WITH NILSEQUENCES

The balance of this paper is dedicated to the proof of Theorem 5.2. Our proof proceeds via the methods from [18] and therefore splits into two tasks. This section accomplishes one them. Recall Definition 5.1 of  $R_i$  for  $1 \leq i \leq r$ . We show here that the function  $R_i$ , when passing to suitable subprogressions and subtracting off its mean value, does not correlate with nilsequences. In Sections 7 and 8 we deal with the second task and construct a pseudorandom majorant for  $R_i$ . To ease notation we shall drop the subscript  $i$  and consider the representation function associated to a typical  $K$  of degree  $n$ .

In order for an arithmetic function to be orthogonal to nilsequences, it first of all needs to be equidistributed in residue classes to small moduli. That is, its average value should not change when passing to subprogressions with respect to small moduli. For this to be valid in our situation, we will choose a product  $W$  of powers of small primes, split

$$R(m) = \sum_{A \pmod{W}} R(m) \mathbf{1}_{m \equiv A \pmod{W}}$$

and consider each of the functions  $m \mapsto R(Wm + A)$  separately. This operation is called the “ $W$ -trick” and was introduced in [17].

Following [27, p.260], let  $w(T) = \log \log T$  and let

$$W = \prod_{p \leq w(T)} p^{\alpha(p)}, \tag{6.1}$$

where  $\alpha(p) = \lceil (C_1 + 1) \log_p \log T \rceil$  for a constant  $1 \leq C_1 \ll 1$  to be specified in Proposition 8.2. In particular,

$$p^{\alpha(p)-1} < (\log T)^{C_1+1} \leq p^{\alpha(p)}.$$

Taking  $T$  sufficiently large, we may henceforth assume that  $M \mid W$ . Moreover, it is clear that  $W = O(T^{o(1)})$ .

Our first result concerns the average order of the  $W$ -tricked functions  $m \mapsto R(Wm + A)$ .

**Lemma 6.1.** *Let  $R = R(m; \mathfrak{X}, \mathbf{b}, M)$  and let  $\epsilon \in \{\pm\}$  be such that  $\mathfrak{X} \cap \mathfrak{D}_+^\epsilon \neq \emptyset$ . For any  $q \in \mathbb{Z}_{>0}$  with  $M \mid q$ , we have*

$$\sum_{\substack{0 < \epsilon m \leq x \\ m \equiv A \pmod{q}}} R(m) = \frac{\varrho(q, A; M)}{q^n} \kappa^\epsilon x + O(qx^{1-1/n}),$$

where  $\varrho(q, A; M)$  is given by (4.3) and (4.4) with  $\mathbf{x}_0 = \mathbf{b}$  and  $\kappa^\epsilon = \kappa^\epsilon(\mathfrak{X})$  is given by (5.2).

*Proof.* Let us write  $\mathfrak{X}(x) = \mathfrak{X} \cap \mathfrak{D}_+^\epsilon(x)$ . Breaking the given sum over  $R(m)$  into residue classes, we find that

$$\sum_{\substack{0 < \epsilon m \leq x \\ m \equiv A \pmod{q}}} R(m) = \sum_{\substack{\mathbf{y} \in (\mathbb{Z}/q\mathbb{Z})^n \\ \mathbf{N}_K(\mathbf{y}) \equiv A \pmod{q} \\ \mathbf{y} \equiv \mathbf{b} \pmod{M}}} \# \{ \mathbf{x} \in \mathbb{Z}^n \cap \mathfrak{X}(x) : \mathbf{x} \equiv \mathbf{y} \pmod{q} \}.$$

The inner cardinality equals  $\#(\mathbb{Z}^n \cap q^{-1}(\mathfrak{X}(x) - \mathbf{y}))$ , which in turn equals

$$\frac{\kappa^\epsilon(\mathfrak{X})x}{q^n} + O(q^{1-n}x^{1-1/n}),$$

by Lemma 3.1. The statement of the lemma easily follows.  $\square$

The results that follow no longer hold for arbitrary residue classes  $A \pmod{W}$  and we will be forced to work with the set of *unexceptional* residue classes

$$\mathcal{A} = \left\{ A \pmod{W} : \begin{array}{l} 0 \leq v_p(A) < v_p(W)/3 \text{ for all } p < w(T) \\ 0 \leq v_p(A) < v_p(M) \text{ for all } p \mid M \\ \varrho(W, A; M) > 0 \end{array} \right\}. \quad (6.2)$$

To justify this, we shall see in Proposition 8.2 that integers that are divisible by a large prime power make a negligible contribution to the asymptotic formula in Theorem 5.2. Consequently, such integers may be excluded from consideration altogether. Next, in view of our assumption that  $p^{v_p(M)} \nmid f_i(\mathbf{a})$  for any  $p \mid M$  and any  $1 \leq i \leq r$ , it is clear that there is no contribution from progressions  $\{m \equiv A \pmod{W}\}$  such that  $v_p(A) \geq v_p(M)$  for any  $p \mid M$ . Finally, when  $\varrho(W, A; M) = 0$  then  $R$  is identically 0 on the progression  $\{m \equiv A \pmod{W}\}$  and so we may exclude these residue classes  $A$  as well.

The next result shows that the function  $m \mapsto R(Wm + A)$  is equidistributed in residue classes to  $w(T)$ -smooth moduli whenever  $A \pmod{W}$  is an unexceptional residue.

**Lemma 6.2.** *Let  $\epsilon \in \{\pm\}$  be such that  $\mathfrak{X} \cap \mathfrak{D}_+^\epsilon \neq \emptyset$ . Let  $A$  be a representative of a class from  $\mathcal{A}$  such that  $0 < \epsilon A < W$ , and let  $a, q \in \mathbb{Z}$  be such that  $0 \leq \epsilon a < q < T/W$ . Suppose further that  $q$  is  $w(T)$ -smooth, and assume that  $T', T'' \in \mathbb{Z}_{>0}$  such that  $T' \asymp T/W$  and  $T'' \asymp T/(qW)$ . Then we have*

$$\mathbb{E}_{0 \leq \epsilon m < T'} R(Wm + A) = \mathbb{E}_{0 \leq \epsilon m < T''} R(W(qm + a) + A) + O(q^2 T^{-1/n+o(1)}).$$

*Proof.* Let  $E_1$  denote the sum on the left hand side and let  $E_2$  be the sum on the right hand side. Since  $T'$  is an integer, the summation range of  $E_1$  may be written as  $0 \leq \epsilon m \leq T' - 1$ .

Since  $m$  and  $A$  are both of sign  $\epsilon$ , we deduce that  $0 < \epsilon(Wm + A) \leq (T' - 1)W + \epsilon A < T'W$ . Thus after a change of variable we have

$$E_1 = \frac{1}{T'} \sum_{\substack{0 < \epsilon m' < T'W \\ m' \equiv A \pmod{W}}} R(m'),$$

and, similarly,

$$E_2 = \frac{1}{T''} \sum_{\substack{0 < \epsilon m' < T''Wq \\ m' \equiv A + Wa \pmod{Wq}}} R(m').$$

Recall that  $M \mid W$ . Two applications of Lemma 6.1 therefore imply that it suffices to prove that

$$\frac{\varrho(W, A; M)}{W^{n-1}} = \frac{\varrho(Wq, A + Wa; M)}{(Wq)^{n-1}}. \quad (6.3)$$

But this follows from the Chinese remainder theorem and applications of Lemma 3.4 for each prime  $p < w(T)$ . Indeed, let  $p < w(T)$ ,  $G = \mathbf{N}_K$ ,  $\ell = v_p(M)$  and let  $m$  be any integer such that  $m \geq v_p(W)$ . Since  $A$  describes an unexceptional residue class, we have  $A \not\equiv 0 \pmod{p^{v_p(W)}}$  and furthermore

$$v_p(M) + v_p(A) + v_p(n) \leq \frac{v_p(W)}{3} + O(1) < \frac{m}{2},$$

provided  $T$  is sufficiently large. Hence, the conditions of Lemma 3.4 are satisfied for large  $T$  and we deduce (6.3) by applying this lemma once for each value of  $m$  in the range  $v_p(W) \leq m < v_p(Wq)$ .  $\square$

The next goal is to establish that the normalised counting function

$$m \mapsto \frac{W^{n-1}}{\varrho(W, A; M)} R(Wm + A)$$

does not correlate with nilsequences if  $A$  is unexceptional. A discussion of the various objects appearing in the following proposition may be found in [27, §§13–15]. A thorough treatment is contained in [19], which is the paper that the results from [27, §§14–16] build on and extend.

**Proposition 6.3.** *Let  $G/\Gamma$  be a nilmanifold of dimension  $m_G \geq 1$ , let  $G_\bullet$  be a filtration of  $G$  of degree  $d \geq 1$ , and let  $g \in \text{poly}(\mathbb{Z}, G_\bullet)$  be a polynomial sequence. Suppose that  $G/\Gamma$  has a  $Q$ -rational Mal'cev basis  $\mathcal{X}$  for some  $Q \geq 2$ , defining a metric  $d_{\mathcal{X}}$  on  $G/\Gamma$ . Suppose that  $F : G/\Gamma \rightarrow [-1, 1]$  is a Lipschitz function. Then for  $\epsilon \in \{\pm\}$ ,  $T' \asymp T/W$  and  $A \in \mathbb{Z}$  with  $A \pmod{W} \in \mathcal{A}$  and  $0 \leq \epsilon A < W$ , we have the estimate*

$$\left| \mathbb{E}_{0 < \epsilon m \leq T'} \left( R(Wm + A) - \frac{\varrho(W, A; M)}{W^{n-1}} \kappa^\epsilon \right) F(g(|m|)\Gamma) \right| \\ \ll_{m_G, d, E} \frac{\varrho(W, A; M)}{W^{n-1}} Q^{O_{m_G, d, E}(1)} \frac{1 + \|F\|_{\text{Lip}}}{(\log \log \log T)^E},$$

for any  $E > 0$ .

Exactly as in [27, Props. 17.1 and 17.2] we deduce the above proposition from a special case involving only “minor arc nilsequences”. This reduction is modelled upon [20, §2] and we will not give the details. The key ingredients are Lemma 6.2 and [27, Thm. 16.4], which is a factorisation theorem for nilsequences. Due to the similar set-up, the choice of parameters from the proof of [27, Prop. 17.1] remains unchanged.

**Proposition 6.4.** *Let  $\epsilon \in \{\pm 1\}$ ,  $T' \asymp T/W$  and  $A \pmod{W} \in \mathcal{A}$  with  $0 \leq \epsilon A < W$ . Suppose that  $\delta \in (0, 1/2)$  and  $S = O(T^{o(1)})$  are parameters such that  $\delta^{-t} \ll_t T$ , for all  $t \in \mathbb{Z}_{>0}$ . Assume that  $(G/\Gamma, d_{\mathcal{X}})$  is an  $m_G$ -dimensional nilmanifold with a filtration  $G_{\bullet}$  of degree  $d$  and that  $g \in \text{poly}(\mathbb{Z}, G_{\bullet})$ . Finally, suppose that for every  $w(T)$ -smooth number  $\tilde{q} \leq S$  the finite sequence  $(g(\tilde{q}m)\Gamma)_{0 < m \leq T'/\tilde{q}}$  is totally  $\delta$ -equidistributed in  $G/\Gamma$ .*

*For every Lipschitz function  $F : G/\Gamma \rightarrow [-1, 1]$  satisfying  $\int_{G/\Gamma} F = 0$ , for every  $w(T)$ -smooth number  $q = O(T^{o(1)})$  and every  $0 \leq b < q$ , for every  $T'' \asymp T/(Wq)$ , there exists  $c \asymp_{m_G, d} 1$  such that*

$$\left| \mathbb{E}_{0 < \epsilon m \leq T''} R(W(qm + b) + A) F(g(|m|)\Gamma) \right| \ll \delta^c (1 + \|F\|_{\text{Lip}}) \frac{\varrho(W, A; M)}{W^{n-1}}.$$

*Proof.* To begin with we note that in the polynomial  $\mathbf{N}_K \in \mathbb{Z}[X_1, \dots, X_n]$  the coefficient of  $X_i^n$  is given by  $N_{K/\mathbb{Q}}(\omega_i) \neq 0$ , for  $1 \leq i \leq n$ .

Our first step is to rewrite the given correlation as a sum over lattice points. A change of variables yields

$$\begin{aligned} & \mathbb{E}_{0 < \epsilon m' \leq T''} R(W(qm' + b) + A) F(g(|m'|)\Gamma) \\ &= \frac{1}{T''} \sum_{\substack{0 < \epsilon m \leq B \\ m \equiv A + Wb \pmod{Wq}}} R(m) F\left(g\left(\frac{m - A - Wb}{\epsilon Wq}\right)\Gamma\right), \end{aligned} \quad (6.4)$$

for some  $B \asymp T$ . Let

$$\mathcal{Y} = \left\{ \mathbf{y} \in (\mathbb{Z}/Wq\mathbb{Z})^n : \begin{array}{l} \mathbf{N}_K(\mathbf{y}) \equiv A + Wb \pmod{Wq} \\ \mathbf{y} \equiv \mathbf{b} \pmod{M} \end{array} \right\},$$

so that  $\#\mathcal{Y} = \varrho(Wq, A + Wb; M)$ . The right hand side of (6.4) becomes

$$\frac{1}{T''} \sum_{\mathbf{y} \in \mathcal{Y}} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ Wq\mathbf{x} + \mathbf{y} \in B^{1/n}\mathfrak{X}(1)}} F\left(g\left(\frac{\mathbf{N}_K(Wq\mathbf{x} + \mathbf{y}) - A - Wb}{\epsilon Wq}\right)\Gamma\right), \quad (6.5)$$

where  $\mathfrak{X}(1) = \{\mathbf{x} \in \mathfrak{X} : 0 < \epsilon \mathbf{N}_K(\mathbf{x}) \leq 1\}$ .

Since the coefficient of  $X_n^n$  in  $\mathbf{N}_K(X_1, \dots, X_n)$  is non-zero, we obtain an integral polynomial of degree  $n$  and leading coefficient  $\epsilon N_{K/\mathbb{Q}}(\omega_n)(Wq)^{n-1}$  when fixing all but the  $n$ th variable in

$$\frac{\mathbf{N}_K(Wq\mathbf{x} + \mathbf{y}) - A - Wb}{\epsilon Wq}.$$

Let  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  denote the projection onto the coordinate plane  $\{x_n = 0\}$ , and let  $P_{\pi(\mathbf{x}), \mathbf{y}}(x) = \gamma_0 + \dots + \gamma_n x^n$  denote the above polynomial, for suitable coefficients  $\gamma_0, \dots, \gamma_n \in \mathbb{Z}$ , with  $\gamma_n = \epsilon N_{K/\mathbb{Q}}(\omega_n)(Wq)^{n-1}$ . If  $Wq\mathbf{x} + \mathbf{y} \in B^{1/n}\mathfrak{X}(1)$ , then it follows that

$\gamma_i \ll B^{(n-i)/n}(Wq)^{i-1}$ , for  $0 \leq i \leq n$ . Thus the hypotheses of [27, Prop. 15.4] are met. We aim to employ this to bound (6.5) by splitting the range of the  $\mathbf{x}$ -summation into lines on which  $\pi(\mathbf{x})$  is constant. With this in mind, we proceed to investigate how such lines intersect the domain  $\mathfrak{X}(1)$ .

We have  $\mathfrak{X}(1) \subset (-\alpha, \alpha)^n$  for some constant  $0 < \alpha = O(1)$ . Let  $\mathbf{a} = (a_1, \dots, a_{n-1})$ , with  $|\mathbf{a}| < \alpha$ , and consider the line  $\ell_{\mathbf{a}} : (-\alpha, \alpha) \rightarrow \mathbb{R}^n$  given by  $\ell_{\mathbf{a}}(x) = (\mathbf{a}, x)$ . For  $\varepsilon \geq 0$ , let  $\partial_\varepsilon \mathfrak{X}(1) \subset \mathbb{R}^n$  denote the set of points at distance at most  $\varepsilon$  to the boundary of the closure of  $\mathfrak{X}(1)$ . We note that the set

$$\{x \in (-\alpha, \alpha) : \ell_{\mathbf{a}}(x) \in \mathfrak{X}(1) \setminus \partial_0 \mathfrak{X}(1)\}$$

is the union of disjoint open intervals. By removing all intervals of length at most  $\varepsilon$ , we obtain a collection of at most  $2\alpha\varepsilon^{-1} \ll \varepsilon^{-1}$  open intervals  $I_1(\mathbf{a}), \dots, I_{k(\mathbf{a})}(\mathbf{a}) \in (-\alpha, \alpha)$  such that any  $x \in (-\alpha, \alpha)$  satisfies the implication

$$\ell_{\mathbf{a}}(x) \in \mathfrak{X}(1) \setminus \partial_\varepsilon \mathfrak{X}(1) \implies x \in I_j(\mathbf{a}) \text{ for some } j \in \{1, \dots, k(\mathbf{a})\}.$$

We will choose a suitable value of  $\varepsilon$  at the end of the proof.

Observe that any interval  $(z_0, z_1) \subset (-\alpha, \alpha)$  can be expressed as a difference of intervals in  $(-\alpha, \alpha)$  that have length at least  $2\alpha/3$ . Indeed,  $z_0$  and  $z_1$  partition  $(-\alpha, \alpha)$  into three (possibly empty) intervals, at least one of which has length at least  $2\alpha/3$ . Thus, one of the three representations

$$(z_0, z_1) = (-\alpha, z_1) \setminus (-\alpha, z_0] = (z_0, \alpha) \setminus [z_1, \alpha)$$

has the required property. For each  $\mathbf{a}$  and  $j \in \{1, \dots, k(\mathbf{a})\}$ , we let  $I_j(\mathbf{a}) = J_j^{(1)}(\mathbf{a}) \setminus J_j^{(2)}(\mathbf{a})$  be such a decomposition, where  $J_j^{(2)}(\mathbf{a})$  is possibly empty.

Abbreviating  $\mathbf{a}' = B^{-1/n}(Wq\mathbf{a} + \pi(\mathbf{y}))$ , we see that (6.5) equals

$$\begin{aligned} & \frac{1}{T''} \sum_{\mathbf{y} \in \mathcal{Y}} \sum_{\substack{\mathbf{a} \in \mathbb{Z}^{n-1} \\ |\mathbf{a}'| < \alpha}} \sum_{j=1}^{k(\mathbf{a})} \sum_{x \in \mathbb{Z}} \left\{ \mathbf{1}_{B^{-1/n}Wqx \in J_j^{(1)}(\mathbf{a}')} - \mathbf{1}_{B^{-1/n}Wqx \in J_j^{(2)}(\mathbf{a}')} \right\} F\left(g(P_{\mathbf{a}, \mathbf{y}}(x))\Gamma\right) \\ & + O\left(\frac{1}{T''} \sum_{\mathbf{y} \in \mathcal{Y}} \#\{\mathbf{x} \in \mathbb{Z}^n : B^{-1/n}(Wq\mathbf{x} + \mathbf{y}) \in \partial_\varepsilon \mathfrak{X}(1)\}\right). \end{aligned} \quad (6.6)$$

Here, the error term accounts for all points in the  $B^{1/n}\varepsilon$ -neighbourhood of the boundary of  $B^{1/n}\mathfrak{X}(1)$ , that were excluded through the choice of intervals  $I_j(\mathbf{a})$ . Observe that we made use of the fact that  $\|F\|_\infty \leq 1$ . Since  $\mathfrak{X}(1)$  is  $(n-1)$ -Lipschitz parametrisable, we have  $\text{vol}(\partial_\varepsilon \mathfrak{X}(1)) \asymp \varepsilon$ . Together with an application of (6.3) this shows that the error term is bounded by

$$T''^{-1} \#\mathcal{Y} \frac{\varepsilon B}{(Wq)^n} \ll \varepsilon B T^{-1} \frac{\varrho(Wq, A + Wb; M)}{(Wq)^{n-1}} \ll \varepsilon \frac{\varrho(W, A; M)}{W^{n-1}}.$$

Turning towards the main term, [27, Prop. 15.4] implies that for every polynomial  $P_{\mathbf{a}, \mathbf{y}}$  there is a  $w(T)$ -smooth integer  $\tilde{q} \in \mathbb{Z}_{>0}$ , with  $\tilde{q} \ll T^{o(1)}$ , and a constant  $c \asymp_{m_G, d} 1$  such



that for each  $0 \leq \tilde{b} < \tilde{q}$  the sequences

$$(g(P_{\mathbf{a}, \mathbf{y}}(\tilde{q}x + \tilde{b}))\Gamma)_{x \leq (T'/\gamma_n \tilde{q}^n)^{1/n}}$$

are totally  $\delta^c$ -equidistributed in  $G/\Gamma$ , provided that  $T$  is large enough. Recall that the leading coefficient of  $P_{\mathbf{a}, \mathbf{y}}$  satisfies  $\gamma_n \asymp (Wq)^{n-1}$ . Since the set

$$\left\{ x \in \mathbb{Z} : \frac{Wq}{B^{1/n}}(\tilde{q}x + \tilde{b}) \in J_j^{(1)}(\mathbf{a}') \right\}$$

is a discrete interval of length

$$\# \left\{ x \in \mathbb{Z} : \frac{Wq}{B^{1/n}}(\tilde{q}x + \tilde{b}) \in J_j^{(1)}(\mathbf{a}') \right\} \asymp \frac{T^{1/n}}{Wq\tilde{q}},$$

we may employ the above total  $\delta^c$ -equidistribution property to deduce that

$$\begin{aligned} \left| \sum_{\substack{x \in \mathbb{Z} \\ B^{-1/n}Wqx \in J_j^{(1)}(\mathbf{a}')}} F(g(P_{\mathbf{a}, \mathbf{y}}(x))\Gamma) \right| &\leq \sum_{\tilde{b}=0}^{\tilde{q}-1} \left| \sum_{\substack{x \in \mathbb{Z} \\ B^{-1/n}Wq(\tilde{q}x + \tilde{b}) \in J_j^{(1)}(\mathbf{a}')}} F(g(P_{\mathbf{a}, \mathbf{y}}(\tilde{q}x + \tilde{b}))\Gamma) \right| \\ &\ll \tilde{q}T^{1/n}(Wq\tilde{q})^{-1}\delta^c\|F\|_{\text{Lip}} \\ &\ll T^{1/n}(Wq)^{-1}\delta^c\|F\|_{\text{Lip}}. \end{aligned}$$

The same holds for  $J_j^{(1)}(\mathbf{a}')$  replaced by any non-empty  $J_j^{(2)}(\mathbf{a}')$ . Hence (6.6) is bounded by

$$\begin{aligned} &\ll T''^{-1} \# \mathcal{Y} \left( \frac{T^{1/n}}{Wq} \right)^{n-1} \varepsilon^{-1} T^{1/n} (Wq)^{-1} \delta^c \|F\|_{\text{Lip}} + \varepsilon \frac{\varrho(W, A; M)}{W^{n-1}} \\ &\ll T''^{-1} \# \mathcal{Y} \frac{T}{(Wq)^n} \varepsilon^{-1} \delta^c \|F\|_{\text{Lip}} + \varepsilon \frac{\varrho(W, A; M)}{W^{n-1}} \\ &\ll \frac{\varrho(W, A; M)}{W^{n-1}} \left( \varepsilon^{-1} \delta^c \|F\|_{\text{Lip}} + \varepsilon \right), \end{aligned}$$

where we applied (6.3). Choosing  $\varepsilon = \delta^{c/2}$  completes the proof.  $\square$

## 7. MAJORANTS FOR POSITIVE MULTIPLICATIVE FUNCTIONS

The aim of this section is to construct for every multiplicative function  $f : \mathbb{Z}_{>0} \rightarrow \mathbb{R}_{>0}$  whose growth is controlled in some precise sense, for every sufficiently small  $\gamma \in (0, 1)$  and for any increasing infinite sequence  $\mathcal{T} = \{T_1 < T_2 < \dots\}$  of sufficiently large positive integers, a family of majorant functions

$$\left( \nu^{(T)} : \{1, \dots, T\} \rightarrow \mathbb{R}_{>0} \right)_{T \in \mathcal{T}}$$

with the following properties:

- (i)  $f(m) \leq C\nu^{(T)}(m)$  for all  $m \leq T$  and some absolute constant  $C > 0$ ;

- (ii)  $\mathbb{E}_{m \leq T} f(m) \asymp \mathbb{E}_{m \leq T} \nu^{(T)}(m)$ ; and
- (iii)  $\nu^{(T)}$  has the structure of a truncated divisor sum. That is to say, it takes the form

$$\nu^{(T)}(m) = \sum_{d \leq T^\gamma} \lambda_d \mathbf{1}_{d|m},$$

for suitable coefficients  $\lambda_d \in \mathbb{R}$ , for all  $m \leq T$  that lie outside a sparse exceptional set.

In [26] such majorant functions were constructed for the divisor function, building on work of Erdős [14]. Shiu [33] observed that Erdős' methods carry over to all multiplicative functions  $f : \mathbb{Z}_{>0} \rightarrow \mathbb{R}_{\geq 0}$  that satisfy the two conditions:

- (a)  $f(p^k) \leq H^k$  for all prime powers; and
- (b)  $f(m) \ll_\delta m^\delta$  as  $m \rightarrow \infty$  for any  $\delta > 0$ .

Equally, the majorant construction from [26, §4] has an analogue for a more general class of multiplicative functions, which we shall describe below. The results in this section do not require condition (b). We employ this condition however in Section 9 when checking the correlation condition. In order to ensure that condition (ii) from above applies to the type of majorant we construct, we impose the further condition that  $g = \mu * f$  is non-negative

**Definition 7.1.** Let  $\mathcal{M}(H)$  denote the set of multiplicative functions  $f : \mathbb{Z}_{>0} \rightarrow \mathbb{R}_{\geq 0}$  such that:

- (a)  $f(p^k) \leq H^k$  for all primes  $p$  and  $k \in \mathbb{Z}_{>0}$ ;
- (b)  $f(m) \ll_\delta m^\delta$  for all  $m \in \mathbb{Z}_{>0}$  and any  $\delta > 0$ ; and
- (c)  $f(p^k) \geq f(p^{k-1})$  for all primes  $p$  and  $k \in \mathbb{Z}_{>0}$ .

Let  $\mathcal{M}'(H)$  denote the set of non-negative multiplicative functions  $f$  satisfying (a) and (b).

Property (c) ensures that any  $f \in \mathcal{M}(H)$  always takes positive values. Moreover, given  $f \in \mathcal{M}(H)$ , we note that  $g = \mu * f$  satisfies  $0 \leq g(p^k) \leq H^k$ .

**Remark 7.2.** Examples of functions which belong to  $\mathcal{M}(H)$ , for suitable  $H$ , include the generalised divisor functions  $\tau_k$ , which appear as Dirichlet coefficients in  $\zeta^k(s)$ , and functions of the form  $h^{\omega(m)}$ , for any real number  $h > 1$ .

For technical reasons we replace all cut-offs, as in (iii) above, by smooth cut-offs. For this purpose, let  $\chi : \mathbb{R} \rightarrow [0, 1]$  be a smooth function that is supported on  $[-1, 1]$ , monotonic on both  $[-1, 0]$  and  $[0, 1]$ , and satisfies  $\chi(x) = 1$  for  $x \in [-1/2, 1/2]$ .

**Definition 7.3** (Truncated multiplicative function). Given a cut-off parameter  $T$  and any multiplicative function  $f : \mathbb{Z}_{>0} \rightarrow \mathbb{R}_{\geq 0}$ , let  $f_\gamma^{(T)} : \{1, \dots, T\} \rightarrow \mathbb{R}_{\geq 0}$  be defined by

$$f_\gamma^{(T)}(m) = \sum_{d \in \mathbb{Z}_{>0}} \mathbf{1}_{d|m} g(d) \chi\left(\frac{\log d}{\log T^\gamma}\right),$$

where  $g = \mu * f$ .

Since  $g$  is non-negative if  $f \in \mathcal{M}(H)$ , we have  $f_\gamma^{(T)}(m) \leq \sum_{d \in \mathbb{Z}_{>0}} \mathbf{1}_{d|m} g(d) = f(m)$  for  $m \leq T$ . The fact that  $\chi(x) = 1$  for  $x \in [-1/2, 1/2]$  implies the lower bound

$$f_\gamma^{(T)}(m) \geq \sum_{d \leq T^{\gamma/2}} \mathbf{1}_{d|m} g(d), \quad (7.1)$$

which is an equality for  $m \leq T^{\gamma/2}$ . The following lemma generalises a result of Erdős, in the form of [26, Lemma 4.1].

**Lemma 7.4.** *Let  $f \in \mathcal{M}(H)$  for  $H > 1$ , let  $C_1 > 1$  be a fixed constant and let  $\xi < 1/2$ . Furthermore, let  $T > 1$  be an arbitrary integer, let  $m \leq T$  and suppose that*

$$f(m) \geq H^\kappa f_{2\xi}^{(T)}(m),$$

*for some  $\kappa > 2/\xi$ . Then one of the following three alternatives holds:*

- (1)  *$m$  is excessively “rough” in the sense that it is divisible by some prime power  $p^a$ ,  $a \geq 2$ , with  $p^a > (\log T)^{C_1}$ ;*
- (2)  *$m$  is excessively “smooth” in the sense that*

$$\prod_{p \leq T^{1/(\log \log T)^3}} p^{v_p(m)} \geq T^{\xi/\log \log T};$$

- (3)  *$m$  has a “cluster” of prime factors in the sense that there is a  $\lambda$  in the interval*

$$\log_2 \kappa - 2 \leq \lambda < \log_2((\log \log T)^3) < 6 \log \log \log T$$

*such that  $m$  has at least  $\xi \kappa(\lambda + 3 - \log_2 \kappa)/100$  distinct prime factors in the superdyadic range  $I_\lambda = [T^{1/2^{\lambda+1}}, T^{1/2^\lambda}]$  and is not divisible by the square of any prime in this range.*

*Proof.* In view of (7.1), the proof of [26, Lemma 4.1] applies when replacing  $\tau_\gamma$  by  $f_{2\xi}^{(T)}$  in such a way that  $\xi$  takes the role of  $\gamma$ . Two obvious changes are necessary:

$$f(m) \leq H^{k-j} f(m') \leq H^{2/\xi} f(m') \leq H^{2/\xi} f_{2\xi}^{(T)}(m) < H^\kappa f_{2\xi}^{(T)}(m),$$

and

$$f(m) \leq H^{a_{j+1} + \dots + a_k} f(m') \leq H^{(r+1)/\xi} f_{2\xi}^{(T)}(m) \leq H^{2r/\xi} f_{2\xi}^{(T)}(m).$$

The rest of the argument is identical. □

The previous lemma allows us to extract majorant functions of truncated divisor sum type, at least outside the following exceptional set.

**Definition 7.5** (Exceptional set). For fixed  $\gamma > 0$ , let  $\mathcal{S} = \mathcal{S}_{C_1, T}$  denote the set of positive integers satisfying condition (1) or (2) of Lemma 7.4 with  $\xi = \gamma/2$ . (This exceptional set is identical to the exceptional set in the divisor function case [26] for  $\gamma/2$  instead of  $\gamma$ .)

Lemma 7.4 will provide us with a majorant function of the correct average order for functions  $f \in \mathcal{M}(H)$ . However, in view of the  $W$ -trick from the previous section, we require majorant functions for each of the functions  $m \mapsto f(Wm + A)$ , where as in (6.1),

$$W = \prod_{p < w(T)} p^{\alpha(p)},$$

with  $w(T) = \log \log T$  and  $\alpha(p) = \lceil (C_1 + 1) \log_p \log T \rceil$ , and where  $0 < A < W$  is such that  $v_p(A) < v_p(W)$  for  $p < w(T)$ . Since  $f(Wm + A) = f(A')f((Wm + A)/A')$ , where  $A' = \gcd(A, W)$ , it suffices to study a function that ignores the contribution from small prime factors. Define the function  $h_T : m \mapsto f(m/m')$ , where

$$m' = \prod_{p < w(T)} p^{v_p(m)},$$

for any integer  $m$ . Then  $h_T$  satisfies (a)–(c) in Definition 7.1 whenever  $f$  does.

**Proposition 7.6** (Majorant). *Let  $\mathcal{T} = \{T_1 < T_2 < \dots\} \subset \mathbb{Z}_{>1}$  be an infinite set of positive integers, and let  $f$  be a non-negative multiplicative function. Assume that for all  $T \in \mathcal{T}$  the function  $h_T : m \mapsto f(m/m')$ , where  $m' = \prod_{p < w(T)} p^{v_p(m)}$ , belongs to  $\mathcal{M}(H)$ .*

*Fix  $\gamma > 0$  of the form  $\gamma = 2^{-z}$  for some  $z \in \mathbb{Z}_{>0}$ . For any positive integers  $\lambda, \kappa$  and  $T > 1$  let*

$$\omega(\lambda, \kappa) = \left\lceil \frac{\gamma \kappa (\lambda + 3 - \log_2 \kappa)}{200} \right\rceil$$

*and  $I_\lambda = [T^{1/2^{\lambda+1}}, T^{1/2^\lambda}]$ , and define the sets*

$$U(\lambda, \kappa) = \begin{cases} \{1\}, & \text{if } \kappa = 4/\gamma \text{ and } \lambda = \log_2 \kappa - 2, \\ \emptyset, & \text{if } \kappa = 4/\gamma \text{ and } \lambda \neq \log_2 \kappa - 2, \\ \left\{ p_1 \dots p_{\omega(\lambda, \kappa)} : \begin{array}{l} p_i \in I_\lambda \text{ distinct primes} \\ f(p_i) \neq 1 \end{array} \right\}, & \text{if } \kappa > 4/\gamma. \end{cases}$$

*Let the family of functions*

$$\left( \nu_f^{(T)} : \{1, \dots, T\} \rightarrow \mathbb{R}_{\geq 0} \right)_{T \in \mathcal{T}}$$

*be defined via*

$$\nu_f^{(T)}(m) = \sum_{\kappa=4/\gamma}^{[(\log \log T)^3]} \sum_{\lambda=\lceil \log_2 \kappa - 2 \rceil}^{\lceil \log_2 ((\log \log T)^3) \rceil} \sum_{u \in U(\lambda, \kappa)} H^\kappa \mathbf{1}_{u|m} f(u) h_\gamma^{(T)} \left( \frac{m}{\prod_{p|u} p^{v_p(m)}} \right) + \mathbf{1}_{m \in \mathcal{S}} h_T(m),$$

*where  $h_\gamma^{(T)}$  is associated to  $h_T$  via Definition 7.3 and where  $\mathcal{S} = \mathcal{S}_{C_1, T}$  is the exceptional set from Definition 7.5.*

*Then, for all sufficiently large  $T > 1$  and for all  $m \leq T$ , we have the majorisation property*

$$f(m) \leq H f(m') \nu_f^{(T)}(m) \ll f(m') \nu_f^{(T)}(m).$$

Furthermore, for any  $0 < A < W$  such that  $v_p(A) < v_p(W)$  for  $p < w(T)$ , we have

$$\mathbb{E}_{m \leq (T-A)/W} \nu_f^{(T)}(Wm + A) \ll \frac{\mathbb{E}_{m \leq (T-A)/W} f(Wm + A)}{f(\gcd(A, W))}, \quad (7.2)$$

as  $T \rightarrow \infty$  through  $\mathcal{T}$ . The implied constant may depend on  $\gamma$  and  $H$ , but not on  $T$ .

**Remark 7.7.** Apart from the term  $\mathbf{1}_{m \in \mathcal{S}} h_T(m)$ , the majorant  $\nu_f^{(T)}$  has a truncated divisor sum structure, since each  $u \in U(\lambda, \kappa)$  satisfies  $u \leq T^\gamma$ , by remark (3) after [26, Prop. 4.2].

**Remark 7.8.** In view of Definition 7.3 it is clear that  $h_\gamma^{(T)}$  is a divisor sum. It is not difficult to deduce some information on the set of positive integers  $d$  that *cannot* occur in this sum. This is the set of integers  $d$  such that  $g(d) = \mu * h_T(d) = 0$ . The definition of  $h_T$  implies that  $g(d) = 0$  whenever  $d$  has a prime factor that is smaller than  $w(T)$ . Similarly,  $g(d) = 0$  if  $d$  has a prime factor  $p > w(T)$ , such that  $h_T(p^{v_p(d)}) = 1$ . The latter condition certainly holds when  $f(p^k) = 1$  for all  $k \in \mathbb{Z}_{>0}$ . Thus the truncated divisor sum  $h_\gamma^{(T)}$  only runs through divisors that are free from prime factors of both these types. Moreover, we remark that the definition of  $U(\lambda, \kappa)$  shows that the sum over  $u$  in  $\nu_f^{(T)}$  only contains divisors that are free from primes  $p$  with  $f(p) = 1$ .

Our proof of Proposition 7.6 does not actually require property (b) of Definition 7.1. This property will be used when establishing the pseudorandomness of our majorant function in Section 9.

*Proof of Proposition 7.6.* We define for each “level”  $\kappa \geq 0$  an exceptional set

$$S(\kappa) = \{m \leq T : h_T(m) \geq H^\kappa h_\gamma^{(T)}(m)\}.$$

For any  $m \leq T$  and any  $\kappa_0 \geq 0$ , we then either have  $m \notin S(\kappa_0)$ , or else there is some integer  $\kappa \geq \kappa_0$  such that  $m \in S(\kappa) \setminus S(\kappa + 1)$ . For  $\kappa_0 = 4/\gamma + 1$ , this yields

$$h_T(m) \leq H^{4/\gamma+1} h_\gamma^{(T)}(m) + \sum_{\kappa > 4/\gamma} H^{\kappa+1} \mathbf{1}_{S(\kappa)}(m) h_\gamma^{(T)}(m). \quad (7.3)$$

We claim that Lemma 7.4, applied with  $\xi = \gamma/2$ , provides an upper bound of the form

$$\mathbf{1}_{S(\kappa)}(m) \leq \sum_{\lambda = \lceil \log_2 \kappa - 2 \rceil}^{\lceil \log_2((\log \log T)^3) \rceil} \sum_{u \in U(\lambda, \kappa)} \mathbf{1}_{u|m},$$

valid for every  $m \notin \mathcal{S}$ .

Taking this claim on trust for the moment, let us first deduce that  $h_T(m) \leq H \nu_f^{(T)}(m)$ . Since  $H > 1$ , this bound is only non-trivial if  $m \notin \mathcal{S}$ . Note that if  $m \in S(\kappa)$  for some  $\kappa > (\log \log T)^3$ , then the third alternative from Lemma 7.4 is empty and we must have  $m \in \mathcal{S}$ . Thus, if  $m \notin \mathcal{S}$ , we can truncate the summation in (7.3) at  $\kappa = \lceil (\log \log T)^3 \rceil$ . Further, if  $m \notin \mathcal{S}$  and  $u|m$  for some  $u \in U(\lambda, \kappa)$ ,  $u > 1$ , then  $v_p(m) = 1$  for any  $p|u$ . Thus, for any  $m \notin \mathcal{S}$ , the properties of  $\chi$  and Definition 7.3 imply

$$\mathbf{1}_{u|m} h_\gamma^{(T)}(m) \leq \mathbf{1}_{u|m} h_T \left( \prod_{p|u} p^{v_p(m)} \right) h_\gamma^{(T)} \left( \frac{m}{\prod_{p|u} p^{v_p(m)}} \right) = \mathbf{1}_{u|m} h_T(u) h_\gamma^{(T)} \left( \frac{m}{\prod_{p|u} p^{v_p(m)}} \right).$$

Inserting the claimed bound on  $\mathbf{1}_{S(\kappa)}(m)$  in all remaining terms of the sum in (7.3), making use of the inequality above, and comparing with the definition of  $\nu_f^{(T)}(m)$ , we indeed obtain that  $h_T(m) \leq H\nu_f^{(T)}(m)$ , provided the summations in  $\kappa$  and  $\lambda$  in  $\nu_f^{(T)}$  contain each at least one term; i.e. provided  $T$  is sufficiently large.

To prove the claim it suffices to check that Lemma 7.4 guarantees for  $m \notin \mathcal{S}$  (that is, for  $m$  which do not have property (1) or (2)) that there actually is a cluster of prime divisors all satisfying  $f(p) \neq 1$ . Thus, suppose  $m \notin \mathcal{S}$ . Every prime  $p$  that can appear in a cluster satisfies  $p > T^{(\log \log T)^{-3/2}}$ , which is larger than both  $w(T)$  and  $(\log T)^{C_1}$ , provided  $T$  is sufficiently large. Hence  $f(p) = h_T(p)$  for such primes and, furthermore,  $p^2 \nmid m$  since  $m \notin \mathcal{S}_{C_1, T}$ . Let

$$q = \prod_{\substack{p > T^{(\log \log T)^{-3/2}} \\ f(p)=1, p|m}} p.$$

Then  $g(d) = 0$  for all  $d|m$  with  $\gcd(d, q) > 1$ , by Remark 7.8. This, in turn, implies that  $h_\gamma^{(T)}(m) = h_\gamma^{(T)}(m/q)$ . Since  $h_T(q) = 1$ , we also have  $h_T(m) = h_T(m/q)$ . Hence  $m/q \in S(\kappa)$  and we may apply Lemma 7.4 to  $m/q$  in order to obtain a cluster of prime factors as required.

It remains to check (7.2). We certainly have

$$\mathbb{E}_{m \leq (T-A)/W} \mathbf{1}_{Wm+A \in \mathcal{S}} h_T(Wm+A) \leq \frac{\mathbb{E}_{m \leq (T-A)/W} f(Wm+A)}{f(\gcd(A, W))},$$

which reduces matters to considering the triple sum from  $\nu_f^{(T)}$ . Since  $2^\lambda \leq (\log \log T)^3$ , any prime divisor  $p$  of an element  $u \in U(\lambda, \kappa)$  satisfies  $p \geq T^{1/(2(\log \log T)^3)}$ , which is larger than  $w(T)$  when  $T$  is large enough. Thus we may assume  $\gcd(u, W) = 1$ . Let  $g = \mu * h_T$ , which is a non-negative multiplicative function. If  $P = \{m \leq T : m \equiv A \pmod{W}\}$ , then

$$\frac{1}{|P|} \sum_{m \in P} \sum_{u \in U(\lambda, \kappa)} \mathbf{1}_{u|m} h_T(u) h_\gamma^{(T)} \left( \frac{m}{\prod_{p|u} p^{v_p(m)}} \right) \leq \frac{2W}{T} \sum_{\substack{m \leq T/A' \\ A'm \in P}} \sum_{u \in U(\lambda, \kappa)} h_T(u) \sum_{\substack{d \leq T^\gamma \\ \gcd(d, u)=1}} g(d) \mathbf{1}_{du|m},$$

where  $A' = \gcd(A, W)$ . Note that  $g(d) = 0$  unless  $\gcd(d, W) = 1$ , by Remark 7.8. Since  $\gcd(du, W) = 1$ , the right hand side above is

$$\begin{aligned} &\ll \sum_{u \in U(\lambda, \kappa)} \frac{h_T(u)}{u} \sum_{\substack{d \leq T^\gamma \\ \gcd(d, uW)=1}} \frac{g(d)}{d} \\ &\ll \sum_{u \in U(\lambda, \kappa)} \frac{H^{\omega(u)}}{u} \sum_{\substack{d \leq T^\gamma \\ \gcd(d, W)=1}} \frac{g(d)}{d}. \end{aligned} \tag{7.4}$$

Note that

$$\sum_{u \in U(\lambda, \kappa)} \frac{H^{\omega(u)}}{u} \ll \frac{H^{\omega(\lambda, \kappa)}}{\omega(\lambda, \kappa)!} \left( \sum_{p \in I_\lambda} \frac{1}{p} \right)^{\omega(\lambda, \kappa)} \ll \frac{(H \log 2 + o(1))^{\omega(\lambda, \kappa)}}{\omega(\lambda, \kappa)!}.$$

If  $\gcd(d, W) = 1$  and  $d \leq T^\gamma$ , then the number of integers  $x \leq T$  for which  $x \equiv A \pmod{W}$  and  $d \mid x$  has order  $T/(Wd)$ . Since  $h_T = 1 * g$ , we deduce that

$$\sum_{\substack{d \leq T^\gamma \\ \gcd(d, W) = 1}} \frac{g(d)}{d} \asymp \frac{W}{T} \sum_{\substack{m \leq T \\ m \equiv A \pmod{W}}} h_T(m). \quad (7.5)$$

Hence the inner sum from (7.4) is bounded by  $T^{-1}W \sum_{m \in P} h_T(m)$ .

The above estimates allow us to bound the average value of  $\nu_f^{(T)}(m) - \mathbf{1}_{m \in \mathcal{S}} h_T(m)$  via

$$\begin{aligned} & \left( \sum_{m \in P} h_T(m) \right)^{-1} \sum_{m \in P} \left( \nu_f^{(T)}(m) - \mathbf{1}_{m \in \mathcal{S}} h_T(m) \right) \\ &= \left( \sum_{m \in P} h_T(m) \right)^{-1} \sum_{m \in P} \sum_{\kappa \geq 4/\gamma} \sum_{\lambda \geq \log_2 \kappa - 2} \sum_{u \in U(\lambda, \kappa)} H^\kappa \mathbf{1}_{u|m} h_\gamma^{(T)}(m) \\ &\ll \sum_{\kappa \geq 4/\gamma} \sum_{\lambda \geq \log_2 \kappa - 2} \frac{H^\kappa \cdot (H \log 2 + o(1))^{\omega(\lambda, \kappa)}}{\omega(\lambda, \kappa)!} \\ &\ll \sum_{\kappa \geq 4/\gamma} \sum_{j \geq 1} H^\kappa \cdot \left( \frac{200 \cdot e \cdot (H \log 2 + o(1))}{\gamma \kappa j} \right)^{\gamma \kappa j / 200} \\ &\ll \sum_{\kappa \geq 4/\gamma} \frac{H^\kappa}{\kappa^{\kappa \gamma / 200}} \left( \sum_{j \geq 1} \left( \frac{200 \cdot e \cdot (H \log 2 + o(1))}{\gamma j} \right)^{\gamma j / 200} \right)^\kappa. \end{aligned}$$

This converges absolutely, and hence completes the proof.  $\square$

Our final objective in this section is to show that the exceptional set  $\mathcal{S}_{C_1, T}$  is negligible when evaluating correlations such as the counting function  $N(T)$  given by (5.1), provided  $C_1$  is sufficiently large.

Let  $\Psi = (\psi_1, \dots, \psi_r) : \mathbb{Z}^s \rightarrow \mathbb{Z}^r$  be a system of non-constant linear polynomials whose non-constant parts are pairwise non-proportional and have coefficients bounded by  $L$  in absolute value. Let  $\mathfrak{K} \subset [-1, 1]^s$  be such that  $\Psi(T\mathfrak{K}) \subset (0, T]^r$  and assume that

$$\mathbb{E}_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \mathbf{1}_{d|\psi_i(\mathbf{m})} \ll_L \frac{1}{d},$$

for any  $d \in \mathbb{Z}_{>0}$  and each  $1 \leq i \leq r$ . Since no  $\psi_i$  is constant, the latter condition is guaranteed to hold when  $\mathfrak{K}$  is convex or when  $\mathfrak{K}$  has an  $(s-1)$ -Lipschitz parametrisable boundary. Then it follows from [14] (cf. [26, Lemmas 3.2 and 3.3]) that the exceptional

set satisfies

$$\mathbb{E}_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \mathbf{1}_{\psi_i(\mathbf{m}) \in \mathcal{S}_{C_1, T}} \ll_{L, C_1} (\log T)^{-C_1/2}, \quad (7.6)$$

for each  $1 \leq i \leq r$ . We shall combine this estimate with the following bound on the  $k$ th moment of a non-negative multiplicative function  $f$ .

**Lemma 7.9** ( $k$ th moment bound for  $f$ ). *Suppose  $f$  satisfies Definition 7.1(a) and (b). Let  $k$  be a positive integer and let  $\Psi$  and  $\mathfrak{K}$  be as above. Then*

$$\mathbb{E}_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \prod_{i=1}^r f^k(\psi_i(\mathbf{m})) \ll_{L, r, k} (\log T)^{O_{r, k, H}(1)}.$$

*Proof.* Let  $g = \mu * f$ . Then Hölder's inequality implies

$$\begin{aligned} \mathbb{E}_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \prod_{i=1}^r f^k(\psi_i(\mathbf{m})) &\leq \prod_{i=1}^r \left( \mathbb{E}_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} f^{kr}(\psi_i(\mathbf{m})) \right)^{1/r} \\ &\leq \prod_{i=1}^r \left( \sum_{d_1, \dots, d_{rk} \leq T} \mathbb{E}_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \prod_{j=1}^{rk} \mathbf{1}_{d_j | \psi_i(\mathbf{m})} g(d_j) \right)^{1/r} \\ &\ll_{L, \varepsilon} \sum_{d_1, \dots, d_{rk} \leq T} \frac{\min\{(d_1 \dots d_{rk})^\varepsilon, H^{\Omega(d_1 \dots d_{rk})}\}}{\text{lcm}(d_1, \dots, d_{rk})} \\ &\ll_{L, \varepsilon} \prod_{p \leq T} \left( 1 + \sum_{\substack{\delta_1, \dots, \delta_{rk} \geq 0 \\ \delta_1 + \dots + \delta_{rk} \geq 1}} \frac{\min\{p^{\varepsilon(\delta_1 + \dots + \delta_{rk})}, H^{\delta_1 + \dots + \delta_{rk}}\}}{p^{\max\{\delta_1, \dots, \delta_{rk}\}}} \right), \end{aligned}$$

for any  $\varepsilon > 0$ . For given  $\ell \in \mathbb{Z}_{>0}$  there are at most  $(1 + \ell)^{rk}$  choices of  $\delta_1, \dots, \delta_{rk}$  for which  $\max\{\delta_1, \dots, \delta_{rk}\} = \ell$ . Taking  $\varepsilon = \frac{1}{2rk}$ , it follows that the right hand side is

$$\begin{aligned} &\leq \prod_{p \leq T} \left( 1 + \sum_{\ell \geq 1} \frac{(1 + \ell)^{rk} \min\{p^{\ell/2}, H^{\ell rk}\}}{p^\ell} \right) \\ &\leq \prod_{p \leq H^{2rk}} \left( 1 + \sum_{\ell \geq 1} \frac{(1 + \ell)^{rk}}{p^{\ell/2}} \right) \prod_{H^{2rk} < p \leq T} \left( 1 + \sum_{\ell \geq 1} \frac{(1 + \ell)^{rk} H^{\ell rk}}{p^\ell} \right) \\ &\ll_{r, k, H} (\log T)^{O_{r, k, H}(1)}, \end{aligned}$$

as required.  $\square$

**Proposition 7.10** (Reduction to unexceptional residues). *Let  $\mathcal{S}_{C_1, T}$  be the exceptional set from Definition 7.5. Suppose that  $f_1, \dots, f_r : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}$  are functions that are all bounded pointwise in modulus by some function  $f \in \mathcal{M}'(H)$ . Let  $i \in \{1, \dots, r\}$ . Suppose that  $f'_i : \{1, \dots, T\} \rightarrow \mathbb{R}$  denotes a function which agrees with  $f_i$  on  $\{1, \dots, T\} \setminus \mathcal{S}_{C_1, T}$  and satisfies  $|f'_i(m)| \leq f(m)$  for all  $m \in \mathcal{S}_{C_1, T}$ . If the parameter  $C_1$  of the exceptional set is*



sufficiently large depending on  $r$  and  $H$ , and if  $\Psi$  and  $\mathfrak{K}$  are as above, then

$$\mathcal{E} = \left| \sum_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \prod_{i=1}^r f_i(\psi_i(\mathbf{m})) - \sum_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \prod_{i=1}^r f'_i(\psi_i(\mathbf{m})) \right| = O(T^s (\log T)^{-C_1/4}).$$

*Proof.* Since

$$\mathcal{E} \leq 2 \sum_{\mathbf{m} \in \mathbb{Z}^s \cap T\mathfrak{K}} \sum_{i=1}^r \mathbf{1}_{\psi_i(\mathbf{m}) \in \mathcal{S}_{C_1, T}} \prod_{j=1}^r f(\psi_j(\mathbf{m})),$$

the proposition follows by the Cauchy–Schwarz inequality from (7.6) and Lemma 7.9.  $\square$

## 8. CONSTRUCTION OF THE MAJORANT

The previous section described the construction of majorants for a general class of positive multiplicative functions. Returning to the proof of Theorem 5.2, we shall now consider the representation functions  $R_i$  from Definition 5.1. Building on the results from Section 7, we construct for each of these representation functions a family of majorant functions

$$(\nu_{R_i}^{(T)} : \{1, \dots, T\} \rightarrow \mathbb{R}_{\geq 0})_{T \in \mathcal{T}}$$

with the properties (i)–(iii) described at the start of Section 7. (The cut-off parameter  $T \in \mathcal{T}$  will later correspond to the parameter  $T$  that appears in Theorem 5.2.)

We begin with an easy estimate for  $R_i$  that relates it to the multiplicative function  $r_{K_i}(|m|)$  whose values are given by the coefficients of the Dedekind zeta function (2.6) for the number field  $K_i/\mathbb{Q}$  of degree  $n_i$ .

**Lemma 8.1.** *We have  $R_i(m) \leq 2|\mu_{K_i}|r_{K_i}(|m|)$ , for non-zero  $m \in \mathbb{Z}$ .*

*Proof.* Replacing  $\mathfrak{X}_i \subset \mathfrak{D}_{i,+}$  by  $\mathfrak{D}_{i,+}$  and dropping the congruence condition from Definition 5.1, we obtain the upper bound

$$R_i(m) \leq \# \{ \mathbf{x} \in \mathbb{Z}^{n_i} \cap \mathfrak{D}_{i,+} : \mathbf{N}_{K_i}(\mathbf{x}) = m \}.$$

The discussion of the unit groups  $U_{K_i}$  and  $U_{K_i}^{(+)}$  in Section 2.2 showed that the index of  $Y_{K_i}^{(+)}$  in  $Y_{K_i}$  is at most two, which implies that

$$\begin{aligned} R_i(m) &\leq 2\# \{ \mathbf{x} \in \mathbb{Z}^{n_i} \cap \mathfrak{D}_i : \mathbf{N}_{K_i}(\mathbf{x}) = m \} \\ &\leq 2|\mu_{K_i}| \# \{ \mathbf{x} \in \mathbb{Z}^{n_i} / U_{K_i} : |\mathbf{N}_{K_i}(\mathbf{x})| = |m| \}, \end{aligned}$$

where  $\mathfrak{D}_i$  is a fundamental domain (in the coordinate space) for the free part of  $U_{K_i}$ . The cardinality in the final line equals  $\# \{ (\alpha) \subset \mathfrak{o} : \mathbf{N}(\alpha) = |m| \}$ , which itself is bounded by  $r_{K_i}(|m|)$ , as required.  $\square$

Let  $\mathcal{S}_{C_1, T}$  be the set from Definition 7.5 and recall the definition (5.1) of  $N(T)$ .

**Proposition 8.2.** *For each  $1 \leq i \leq r$ , let  $R'_i : \{-T, \dots, T\} \rightarrow \mathbb{R}_{\geq 0}$  denote a function such that  $R'_i(m) = R_i(m)$  for all  $m$  satisfying  $|m| \in \{1, \dots, T\} \setminus \mathcal{S}_{C_1, T}$ , and which further satisfies  $0 \leq R'_i(m) \leq R_i(m)$  when  $|m| \in \mathcal{S}_{C_1, T}$ . If the parameter  $C_1$  of the exceptional set is sufficiently large, then*

$$N(T) = \beta_\infty \prod_p \beta_p \cdot T^s + o(T^s)$$

if and only if

$$\sum_{\substack{\mathbf{u} \in \mathbb{Z}^s \cap T\mathfrak{K} \\ \mathbf{u} \equiv \mathbf{a} \pmod{M}}} \prod_{i=1}^r R'_i(f_i(\mathbf{u})) = \beta_\infty \prod_p \beta_p \cdot T^s + o(T^s).$$

*Proof.* We have  $R_i(m) \leq 2|\mu_{K_i}|\tau(|m|)^{n_i}$  by Lemma 8.1 and (2.10), where  $\tau^{n_i} \in \mathcal{M}(2^{n_i})$ . For  $\epsilon = (\epsilon_1, \dots, \epsilon_r) \in \{\pm 1\}^r$  we let  $\mathfrak{K}_\epsilon = \mathfrak{K} \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r})$ . Then we may decompose  $\mathfrak{K}$  as the union of the  $2^r$  sets  $\mathfrak{K}_\epsilon$ , together with one set  $\mathfrak{K}_0$  such that at each point  $\mathbf{u} \in \mathfrak{K}_0$  at least one  $f_i$  vanishes. We may discard  $\mathfrak{K}_0$  since  $\prod_{i=1}^r R_i(f_i(\mathbf{u})) = 0$  at all of its integral points. Proposition 7.10 may be applied separately to each of the  $2^r$  remaining sums, by reinterpreting  $R_i$  and  $R'_i$  as functions on  $\{1, \dots, T\}$  via  $m \mapsto R_i(\epsilon_i m)$ .  $\square$

Proposition 8.2 allows us to work with functions

$$R'_i(m) = \begin{cases} R_i(m), & \text{if } |m| \notin \mathcal{S}_{C_1, T}, \\ 0, & \text{if } |m| \in \mathcal{S}_{C_1, T}, \end{cases} \quad (8.1)$$

instead of the original counting functions. Thus the majorant function only needs to majorise  $R_i$  outside the exceptional set.

**Remark 8.3.** For later applications it is convenient at this point to note that

$$\mathbb{E}_{|m| < T/W} R_i(Wm + A) \mathbf{1}_{Wm+A \in \mathcal{S}_{C_1, T}} \ll (\log T)^{-C_1/4} \quad (8.2)$$

and

$$\begin{aligned} \mathbb{E}_{m \leq (T-A)/W} R'_i(Wm + A) &= \mathbb{E}_{m \leq (T-A)/W} R_i(Wm + A) + O((\log T)^{-C_1/4}) \\ &\asymp \frac{\varrho_i(W, A; M)}{W^{n_i-1}}. \end{aligned} \quad (8.3)$$

The first bound follows from Proposition 7.10, applied with (cf. Proposition 8.2)  $r = s = 1$ ,  $f_1 = R_i$  and  $\psi_1(m) = m$ . The second part follows by combining (8.2) with Lemma 6.1.

For most of the remainder of this section we consider a typical representation function and will drop the index  $i$  in such situations. Lemma 8.1 implies that  $R(m) \ll r_K(|m|)$ . Taking  $q = 1$  in Lemma 6.1, we deduce that for  $\epsilon \in \{\pm 1\}$

$$\frac{1}{T} \sum_{0 < \epsilon m \leq T} R(m) \sim \kappa^\epsilon, \quad (T \rightarrow \infty),$$

whereas (2.8) yields

$$\frac{1}{T} \sum_{0 < m \leq T} r_K(m) \sim h\kappa, \quad (T \rightarrow \infty).$$

Hence  $R$  is majorised by  $r_K$  and has the same average order as it.

Given  $T \in \mathbb{Z}_{>0}$ , consider  $P_A = \{m \equiv A \pmod{W}\}$ , for  $A \in \mathcal{A}$  given by (6.2) and  $W$  as in (6.1). Lemma 4.2 only provides us with precise information on  $\varrho(p^m, A)$  when  $p \nmid D_K$  and  $v_p(A) < m$ . This limits our ability to deduce that  $R$  and  $r_K$  have the same average order on progressions  $P_A$  unless  $\gcd(A, D_K) = 1$ . For this reason we will need to refine the bound on  $R$  from Lemma 8.1 to one that is tighter at integers  $m$  with  $\gcd(m, D_K) > 1$ . We set

$$W_0 = \prod_{p|D_K} p^{\alpha(p)}$$

and proceed to establish the following result.

**Lemma 8.4.** *Let  $A \in \mathcal{A}$  and write  $A_0 = \gcd(A, W_0)$ . Then*

$$R(Wm + A) \ll \frac{\varrho(W_0, A)}{W_0^{n-1}} r_K \left( \frac{|Wm + A|}{A_0} \right).$$

*Proof.* The following shorter proof of this result was suggested to us by the referee. To begin with, Lemma 8.1 implies that  $R(Wm + A) \ll r_K(|Wm + A|)$ . Moreover, for any  $A \in \mathcal{A}$  it is clear that  $A_0$  is coprime to  $|Wm + A|/A_0$ . Since  $r_K$  is multiplicative it is therefore enough to show that

$$r_K(A_0) \ll \frac{\varrho(W_0, A)}{W_0^{n-1}}.$$

Now since  $A \in \mathcal{A}$  we must have  $v_p(A_0) < v_p(M) \ll 1$  for any  $p \mid W_0$ . Thus  $r_K(A_0) \ll 1$  by (2.10). Hence it remains to prove that  $\varrho(W_0, A)/W_0^{n-1} \gg 1$ . Let  $p \mid D_K$ . Setting

$$m_0 = 1 + 2(v_p(M) + v_p(A) + v_p(n)) < 1 + 4v_p(M) + 2v_p(n) \ll 1,$$

we apply Lemma 3.4 to obtain

$$\frac{\varrho(p^{\alpha(p)}, A)}{p^{\alpha(p)(n-1)}} \geq \frac{\varrho(p^{\alpha(p)}, A; p^{v_p(M)})}{p^{\alpha(p)(n-1)}} = \frac{\varrho(p^{m_0}, A; p^{v_p(M)})}{p^{m_0(n-1)}} \geq \frac{1}{p^{m_0(n-1)}} \gg 1, \quad (8.4)$$

since  $\varrho(W, A; M) > 0$  in (6.2). This completes the proof of the lemma.  $\square$

In view of Lemma 8.4 we are therefore led to construct a majorant function for each restriction of the multiplicative function  $r_K$  to a progression  $P_A$ , with  $A \in \mathcal{A}$ . Given  $T \in \mathbb{Z}_{>1}$ , we write

$$m' = \prod_{p < w(T)} p^{v_p(m)}, \quad (8.5)$$

for  $m \in \mathbb{Z}$ . Then  $r_K(m')$  is constant for  $m \in P_A$ . Thus we seek a majorant for the function  $m \mapsto r_K(m/m')$ , which is free from the contributions of small prime factors.

Note that  $r_K$  is an unbounded function with sparse support. Our next aim, accomplished in Lemma 8.5 below, is to simplify the task by separating these two properties, replacing  $r_K$  by the product of a bounded function with sparse support and an unbounded function with dense support.

In general we write  $\langle \mathcal{P} \rangle = \{n \in \mathbb{Z}_{>0} : p \mid n \Rightarrow p \in \mathcal{P}\}$  for a set  $\mathcal{P}$  of rational primes. Recall the definition (2.12) of  $\mathcal{P}_0, \mathcal{P}_1$  and  $\mathcal{P}_2$ . Additionally, we require the sets

$$\overline{\mathcal{P}}_j = \{p \in \mathcal{P}_j : p > w(T)\},$$

for  $j = 1, 2$ . It follows from (2.11) that the restriction of  $r_K$  to square-free numbers is supported on  $\langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle$ . Let  $r_{\text{res}}$  denote the multiplicative function defined via

$$r_{\text{res}}(p^m) = \begin{cases} r_K(p^m), & \text{if } p \in \mathcal{P}_0 \cup \mathcal{P}_1, \\ 1, & \text{if } p \in \mathcal{P}_2. \end{cases} \quad (8.6)$$

We have the following result.

**Lemma 8.5.** *For all  $m \in \mathbb{Z}_{>0}$ , we have*

$$r_K(m) \leq r_{\text{res}}(m) \sum_{\substack{q \in \langle \mathcal{P}_2 \rangle \\ v_p(q) \neq 1 \ \forall p}} \mathbf{1}_{q|m} \tau(q)^n \mathbf{1}_{\langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle} \left( \frac{m}{q} \right).$$

*Proof.* If  $r_K(m)$  is positive, then  $m$  has no prime divisor  $p \in \mathcal{P}_2$  for which  $p^2 \nmid m$ . In this case the sum on the right hand side has exactly one term, corresponding to the factorisation of  $m$  into the product of  $q \in \langle \mathcal{P}_2 \rangle$  and  $m/q \in \langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle$ . The multiplicativity of  $r_K$  implies

$$r_K(m) = r_K(q) r_K\left(\frac{m}{q}\right) = r_K(q) r_{\text{res}}\left(\frac{m}{q}\right) = r_K(q) r_{\text{res}}(m) \leq \tau(q)^n r_{\text{res}}(m),$$

where we used (2.10) to bound  $r_K(q)$ . □

As a direct consequence of this lemma, we obtain

$$r_K(m) \leq r_K(m') \cdot r_{\text{res}}\left(\frac{m}{m'}\right) \sum_{\substack{q \in \langle \mathcal{P}_2 \rangle \\ v_p(q) \neq 1 \ \forall p}} \mathbf{1}_{q|m} \tau(q)^n \mathbf{1}_{\langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle} \left( \frac{m}{qm'} \right), \quad (8.7)$$

where  $m'$  is given by (8.5) and we have observed that  $r_K(m) = r_K(m') r_K(m/m')$ , by multiplicativity. In view of (8.7) we proceed by constructing two families of majorant functions in Sections 8.1 and 8.2: one for the positive multiplicative function  $m \mapsto r_{\text{res}}(m/m')$  and one for the characteristic function  $\mathbf{1}_{\langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle}$ . Inserting these majorants into the bound (8.7), we will obtain a family of majorant functions for  $r_K$ . In Section 9 we check that the resulting majorants for  $r_K$  form a family of pseudorandom majorants when restricting them to the arithmetic progressions  $P_A$ .

**8.1. Majorant for  $r_{\text{res}}$ .** Our first task is to check that Proposition 7.6 applies to the function  $r_{\text{res}}$ .

**Lemma 8.6.** *Let  $T \in \mathbb{Z}_{>1}$ . Given  $m \in \mathbb{Z}_{>0}$ , let  $m'$  be defined by (8.5). If  $h_T$  denotes the function  $m \mapsto r_{\text{res}}(m/m')$ , then  $h_T \in \mathcal{M}(2^n)$ , provided that  $T \geq \exp(\exp(2|D_K|))$ .*

*Proof.* We need to check conditions (a)–(c) in Definition 7.1. By (2.10), we have

$$r_{\text{res}}(p^m) \leq r_K(p^m) \leq (m+1)^n = (\tau(p^m))^n \leq 2^{mn}.$$

Thus part (a) holds with  $H = 2^n$ . Part (b) follows immediately from the respective property for the divisor function. To check part (c), we may restrict attention to  $p \in \mathcal{P}_1 \cup \mathcal{P}_2$ , since  $p < w(T)$  when  $p \in \mathcal{P}_0$  and  $T \geq \exp(\exp(2|D_K|))$ . Recalling (8.6) we see that condition (c) is trivially satisfied for  $p \in \mathcal{P}_2$ . If  $p \in \mathcal{P}_1$ , then there is a prime ideal  $\mathfrak{p}|(p)$  of residue degree 1 in  $K/\mathbb{Q}$ . Thus, if  $\mathfrak{a} \subset \mathfrak{o}$  is counted by  $r_{\text{res}}(p^m)$ , that is to say  $N \mathfrak{a} = p^m$ , then  $\mathfrak{a}\mathfrak{p}$  is an ideal counted by  $r_{\text{res}}(p^{m+1})$ . Hence  $r_{\text{res}}(p^m) \leq r_{\text{res}}(p^{m+1})$ , as required for (c).  $\square$

Lemma 8.6 implies that  $r_{i,\text{res}}(m/m') \in \mathcal{M}(2^{n_i})$  for each  $1 \leq i \leq r$ , with  $m'$  given by (8.5). Taking  $f = r_{i,\text{res}}$  and  $h_T = r_{i,\text{res}}(m/m')$ , let  $h_\gamma^{(T)}$  be as in Definition 7.3, with  $g = \mu * h_T$ . Let

$$\nu_i^{(T)}(m) = \sum_{\kappa=4/\gamma}^{[(\log \log T)^3]} \sum_{\lambda=\lceil \log_2 \kappa - 2 \rceil}^{\lceil \log_2((\log \log T)^3) \rceil} \sum_{u \in U(\lambda, \kappa)} 2^{n_i \kappa} \mathbf{1}_{u|m} r_{i,\text{res}}(u) h_\gamma^{(T)} \left( \frac{m}{m' \prod_{p|u} p^{v_p(m)}} \right). \quad (8.8)$$

Then Proposition 7.6 implies that  $\nu_i^{(T)}(m)$  majorises  $r_{i,\text{res}}(m/m')$  on  $\{1, \dots, T\} \setminus \mathcal{S}_{C_1, T}$ . According to Remark 7.8, furthermore, it is a truncated divisor sum that only involves divisors from  $\langle \overline{\mathcal{P}}_1^{(i)} \rangle$  provided  $T \geq \exp(\exp(2|D_{K_i}|))$ . Indeed, if  $p \in \mathcal{P}_0^{(i)}$  then  $p < w(T)$ , and if  $p \in \overline{\mathcal{P}}_2^{(i)}$  then  $r_{i,\text{res}}(p^k) = 1$  for any  $k \in \mathbb{Z}_{>0}$ .

Proposition 7.6 provides the upper bound

$$\mathbb{E}_{m \leq (T-A)/W} \nu_i^{(T)}(Wm + A) \ll \frac{\mathbb{E}_{m \leq (T-A)/W} r_{i,\text{res}}(Wm + A)}{r_{i,\text{res}}(\gcd(A, W))},$$

for any  $A$  belonging to the set  $\mathcal{A}_i$  defined in (6.2). We proceed by deducing the following bound in terms of the arithmetic data that is involved.

**Lemma 8.7.** *For  $A \in \mathcal{A}_i$  we have*

$$\mathbb{E}_{m \leq (T-A)/W} \nu_i^{(T)}(Wm + A) \ll (\log T)^{1-\delta_i} \prod_{\substack{p < w(T) \\ p \in \mathcal{P}_0^{(i)} \cup \mathcal{P}_1^{(i)}}} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{\mathfrak{p}|p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left(1 - \frac{1}{N \mathfrak{p}}\right),$$

where  $\mathcal{P}_0^{(i)}, \mathcal{P}_1^{(i)}, \mathcal{P}_2^{(i)}$  are the sets (2.12) of rational primes corresponding to  $K_i$  and  $\delta_i$  is the Dirichlet density of  $\mathcal{P}_1^{(i)}$ .

*Proof.* Let  $g = \mu * h_T$  and  $h_T = r_{i,\text{res}}(m/m')$  with  $m' = \prod_{p < w(T)} p^{v_p(m)}$ , as before. We may assume that  $T$  is large enough to guarantee that  $g$  is non-negative, with  $g \in \mathcal{M}'(2^{n_i})$ . Then it follows from (7.5) that

$$\frac{\mathbb{E}_{m \leq (T-A)/W} r_{i,\text{res}}(Wm + A)}{r_{i,\text{res}}(\gcd(A, W))} \ll \sum_{d \leq T} \frac{g(d)}{d} \leq \prod_{p \leq T} \left(1 + \sum_{k \geq 1} \frac{g(p^k)}{p^k}\right) \ll \exp \left( \sum_{p \leq T} \frac{g(p)}{p} \right).$$

Since  $r_{i,\text{res}}(p) = r_{K_i}(p) + \mathbf{1}_{\mathcal{P}_2^{(i)}}(p)$  and  $g(p) = 0$  for  $p \leq w(T)$ , the sum in the argument of the exponential function is equal to

$$\sum_{p \leq T} \frac{g(p)}{p} = \sum_{p \leq T} \frac{r_{K_i}(p) - 1}{p} + \sum_{p \leq T} \frac{\mathbf{1}_{\mathcal{P}_2^{(i)}}(p)}{p} + \sum_{\substack{p < w(T) \\ p \in \mathcal{P}_0^{(i)} \cup \mathcal{P}_1^{(i)}}} \left( \frac{1}{p} - \frac{r_{K_i}(p)}{p} \right).$$

The prime ideal theorem [24, Satz 192] implies that the first sum is

$$\sum_{p \leq T} \frac{r_{K_i}(p) - 1}{p} = \log \log T - \log \log T + O(1) = O(1).$$

Corollary 2.4 shows that the second sum satisfies

$$\exp \left( \sum_{p \leq T} \frac{\mathbf{1}_{\mathcal{P}_2^{(i)}}(p)}{p} \right) \asymp (\log T)^{1-\delta_i}.$$

For the final sum we obtain

$$\exp \left( \sum_{\substack{p < w(T) \\ p \in \mathcal{P}_0^{(i)} \cup \mathcal{P}_1^{(i)}}} \left( \frac{1}{p} - \frac{r_{K_i}(p)}{p} \right) \right) \asymp \prod_{\substack{p < w(T) \\ p \in \mathcal{P}_0^{(i)} \cup \mathcal{P}_1^{(i)}}} \left( 1 - \frac{1}{p} \right)^{-1} \prod_{\substack{\mathfrak{p} | p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left( 1 - \frac{1}{N \mathfrak{p}} \right)$$

by combining the approximation  $\log(1 \pm m^{-1}) = \pm m^{-1} + O(m^{-2})$ , valid for integers  $m > 2$ , with the identity  $r_{K_i}(p) = \#\{\mathfrak{p} \mid (p) : \mathfrak{p} \subset \mathfrak{o}_{K_i}, N \mathfrak{p} = p\}$ . This completes the proof.  $\square$

**8.2. Sieve majorant.** In this section we drop the index  $i$  and work with a typical number field  $K$  of degree  $n$  over  $\mathbb{Q}$ . Our next objective is to construct a majorant function of the correct average order for the characteristic function  $\mathbf{1}_{\langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle}$  in any of the arithmetic progressions  $\{m \equiv A \pmod{W}\}$ , for  $A \in \mathcal{A}$ .

Let  $\chi : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  be a smooth even function with  $\text{supp } \chi \subset [-1, 1]$  and  $\chi(x) = 1$  for  $x \in [-1/2, 1/2]$ . As before we let  $\gamma > 0$ , to be viewed as a small fixed constant. In analogy to the construction from [18, App. D], which itself builds on work of Goldston and Yıldırım [15, 16], we consider the functions  $\nu_{\text{sieve}}^{(T)}, \nu'_{\text{sieve}}^{(T)} : \{1, \dots, T\} \rightarrow \mathbb{R}_{\geq 0}$ , defined via

$$\nu_{\text{sieve}}^{(T)}(m) = \left( \sum_{\substack{d \in \langle \mathcal{P}_2 \rangle \\ d | m}} \mu(d) \chi \left( \frac{\log d}{\log T^\gamma} \right) \right)^2 \quad (8.9)$$

and

$$\nu'_{\text{sieve}}^{(T)}(m) = \sum_{\substack{q \in \langle \mathcal{P}_2 \rangle \\ v_p(q) \neq 1 \ \forall p}} \mathbf{1}_{q|m} \tau(q)^n \chi \left( \frac{\log q}{\log T^\gamma} \right) \nu_{\text{sieve}}^{(T)} \left( \frac{m}{q} \right). \quad (8.10)$$

Both of these functions are non-negative. Moreover we note that for  $m \in \langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle$  we have  $\nu_{\text{sieve}}^{(T)}(m) = 1$ . Hence  $\nu_{\text{sieve}}^{(T)}$  majorises  $\mathbf{1}_{\langle \mathcal{P}_0 \cup \mathcal{P}_1 \rangle}$ . The main goal of this section is to establish the following lemma.

**Lemma 8.8.** *For every  $A \in \mathcal{A}$  we have*

$$\frac{W}{T} \sum_{m \leq (T-A)/W} \nu_{\text{sieve}}^{(T)}(Wm + A) \ll (\log T)^{\delta-1} \prod_{\substack{p \in \mathcal{P}_2 \\ p \leq w(T)}} \left(1 + \frac{1}{p}\right),$$

where  $\delta$  is the Dirichlet density of  $\mathcal{P}_1$ .

The key element used for both the proof of Lemma 8.8 and for asymptotically evaluating linear correlations of  $\nu_{\text{sieve}}'$  in the next section is the observation, due to Green and Tao [18, App. D], that one can turn the smooth cut-off  $\chi$  in (8.9) into multiplicative functions as follows. Let  $\vartheta$  be the transform of  $\chi$  that is defined via

$$e^x \chi(x) = \int_{\mathbb{R}} \vartheta(\xi) e^{-ix\xi} d\xi.$$

Recall that  $\chi$  has compact support and is smooth. Fourier inversion and partial integration therefore yield the bound

$$\vartheta(\xi) \ll_E (1 + |\xi|)^{-E} \quad (8.11)$$

for any  $E > 0$ . Following [18, App. D], we make use of this rapid decay to truncate the integral representation of  $\chi$  which will enable us to swap integrations and summations later on. Let  $I = \{\xi \in \mathbb{R} : |\xi| \leq \sqrt{\log T^\gamma}\}$ , then for any  $m \in \mathbb{Z}_{>0}$  we have

$$\begin{aligned} \chi\left(\frac{\log m}{\log T^\gamma}\right) &= \int_{\mathbb{R}} m^{-\frac{1+i\xi}{\log T^\gamma}} \vartheta(\xi) d\xi \\ &= \int_I m^{-\frac{1+i\xi}{\log T^\gamma}} \vartheta(\xi) d\xi + O_E\left(\frac{m^{-1/\log T^\gamma}}{(\log T^\gamma)^E}\right). \end{aligned} \quad (8.12)$$

*Proof of Lemma 8.8.* We begin by estimating, for any parameter  $T' \leq T$  and  $0 \leq A' < W$ , the sum

$$S(T') = \sum_{m \leq T'/W} \nu_{\text{sieve}}^{(T)}(Wm + A').$$

We will show that

$$S(T') \ll_E \frac{T'}{W} (\log T)^{\delta-1} \prod_{\substack{p \in \mathcal{P}_2 \\ p < w(T)}} \left(1 - \frac{1}{p}\right)^{-1} + \frac{T'}{W(\log T)^E} + T^{2\gamma}, \quad (8.13)$$

for any  $E > 0$ . In order to bound the average order of  $\nu_{\text{sieve}}^{(T)}(Wm + A)$  we apply this estimate with  $T' = T/q - A'$ , for  $q \leq T^\gamma$  and for  $A'$  with  $qA' \equiv A \pmod{W}$ . Since

$W \ll T^{o(1)}$ , the first term in the bound dominates and we obtain

$$\begin{aligned} \left( (\log T)^{1-\delta} \prod_{\substack{p \in \mathcal{P}_2 \\ p < w(T)}} \left( 1 - \frac{1}{p} \right) \right) \sum_{m \leq (T-A)/W} \nu'_{\text{sieve}}(T)(Wm + A) &\ll \frac{T}{W} \sum_{\substack{q \in \langle \mathcal{P}_2 \rangle \\ v_p(q) \neq 1 \ \forall p}} \frac{\tau(q)^n}{q} \\ &\ll \frac{T}{W}. \end{aligned}$$

This shows that the lemma will follow if we can establish (8.13).

Let  $d \in \langle \overline{\mathcal{P}_2} \rangle$ . We first note that

$$\# \{m \leq T'/W : Wm + A' \equiv 0 \pmod{d}\} = \frac{T'}{Wd} + O(1),$$

since  $\gcd(d, W) = 1$ . Hence

$$\begin{aligned} S(T') &= \sum_{m \leq T'/W} \sum_{\substack{d, d' \in \langle \overline{\mathcal{P}_2} \rangle \\ [d, d'] | Wm + A'}} \mu(d) \mu(d') \chi\left(\frac{\log d}{\log T^\gamma}\right) \chi\left(\frac{\log d'}{\log T^\gamma}\right) \\ &= \sum_{d, d' \in \langle \overline{\mathcal{P}_2} \rangle} \mu(d) \mu(d') \chi\left(\frac{\log d}{\log T^\gamma}\right) \chi\left(\frac{\log d'}{\log T^\gamma}\right) \left( \frac{T'}{W[d, d']} + O(1) \right). \end{aligned}$$

The overall contribution from the error term is  $O(T^{2\gamma})$ . Applying (8.12), we obtain

$$\begin{aligned} S(T') &= \frac{T'}{W} \sum_{d, d' \in \langle \overline{\mathcal{P}_2} \rangle} \frac{\mu(d) \mu(d')}{[d, d']} \int_I \int_I d^{-\frac{1+i\xi}{\log T^\gamma}} d'^{-\frac{1+i\xi'}{\log T^\gamma}} \vartheta(\xi) \vartheta(\xi') \, d\xi \, d\xi' \\ &\quad + O_E \left( \frac{T'}{W(\log T^\gamma)^E} \sum_{d, d'} \frac{(dd')^{-1/\log T^\gamma}}{[d, d']} \right) + O(T^{2\gamma}), \end{aligned}$$

for any  $E > 0$ . Let us denote the main term, temporarily, by  $M(T')$ . The first of the error terms may be bounded by noting that

$$\begin{aligned} \frac{T'}{W(\log T^\gamma)^E} \sum_{d, d'} \frac{(dd')^{-1/\log T^\gamma}}{[d, d']} &\leq \frac{T'}{W(\log T^\gamma)^E} \sum_{d, d', d''} (dd'd'')^{-1-1/\log T^\gamma} \\ &\leq \frac{T'}{W(\log T^\gamma)^{E-3}}. \end{aligned}$$

Thus both error terms are satisfactory for (8.13), on redefining  $E$ .



It remains to estimate the main term  $M(T')$ . On interchanging the sum over  $d, d'$  with the double integral and taking the Euler product, we obtain

$$\begin{aligned} |M(T')| &\leq \frac{T'}{W} \left| \int_I \int_I \prod_{p \in \overline{\mathcal{P}}_2} \left( 1 - p^{-1 - \frac{1+i\xi}{\log T^\gamma}} - p^{-1 - \frac{1+i\xi'}{\log T^\gamma}} + p^{-1 - \frac{1+i\xi+1+i\xi'}{\log T^\gamma}} \right) \vartheta(\xi) \vartheta(\xi') \, d\xi \, d\xi' \right| \\ &\ll \frac{T'}{W} \int_I \int_I |\Pi(\xi, \xi') \vartheta(\xi) \vartheta(\xi')| \, d\xi \, d\xi', \end{aligned}$$

where

$$\Pi(\xi, \xi') = \prod_{p \in \overline{\mathcal{P}}_2} \left( 1 - p^{-1 - \frac{1+i\xi}{\log T^\gamma}} \right) \left( 1 - p^{-1 - \frac{1+i\xi'}{\log T^\gamma}} \right) \left( 1 + p^{-1 - \frac{1+i\xi+1+i\xi'}{\log T^\gamma}} \right).$$

We denote the final integral by

$$\mathcal{J} = \int_I \int_I |\Pi(\xi, \xi') \vartheta(\xi) \vartheta(\xi')| \, d\xi \, d\xi'.$$

Our aim is to estimate  $\mathcal{J}$  by bounding the product  $\Pi(\xi, \xi')$  from above.

The product  $\Pi(\xi, \xi')$  is intimately related to the Euler product

$$F(s) = \prod_{p \in \overline{\mathcal{P}}_2} \left( 1 - \frac{1}{p^s} \right)^{-1}, \quad (\Re(s) > 1)$$

that we met in (2.13). From Lemma 2.3 we deduce that there is a function  $G(s)$ , which is holomorphic and non-zero in the closed half-plane  $\Re(s) \geq 1$ , such that

$$F(1+s) = G(1+s) \left( \frac{1}{s} + O(1) \right)^{1-\delta} = G(1+s) \frac{1}{s^{1-\delta}} (1 + O(|s|))^{1-\delta}$$

for  $\Re(s) > 0$ . The primes in  $\Pi(\xi, \xi')$  only run over  $\overline{\mathcal{P}}_2$ . Thus, since we may freely disregard finitely many primes, it suffices to ally our knowledge of  $F(s)$  with an investigation of

$$\tilde{F}(s) = \prod_{\substack{p \in \overline{\mathcal{P}}_2 \\ C < p < w(T)}} \left( 1 - \frac{1}{p^s} \right)^{-1}$$

near  $s = 1$ , for a suitable absolute constant  $C = O(1)$ . Lemma 3.2 applies to  $\tilde{F}(s)$  with  $h = \mathbf{1}_{\langle \mathcal{P}_2 \rangle}$  and  $H = 1$ . The primes in  $\tilde{F}$  run up to  $x = w(T) = \log \log T$  and we are interested in  $s_0 = s$  satisfying  $\Re(s) = (\log T^\gamma)^{-1}$  or  $\Re(s) = 2(\log T^\gamma)^{-1}$  and  $|s| \leq 3(\log T^\gamma)^{-1/2} = 3\gamma^{-1/2}e^{-w(T)/2}$ . Thus the conditions of Lemma 3.2 are satisfied and we obtain

$$|\tilde{F}(1+s)| \asymp \tilde{F}(1) \asymp \prod_{\substack{p \in \overline{\mathcal{P}}_2 \\ p < w(T)}} \left( 1 - \frac{1}{p} \right)^{-1}.$$

Thus, invoking (8.11), we obtain

$$\begin{aligned}
\mathcal{J} &\ll \int_I \int_I \left| \frac{1+i\xi}{\log T^\gamma} \right|^{1-\delta} \left| \frac{1+i\xi'}{\log T^\gamma} \right|^{1-\delta} \left| \frac{1+i\xi+1+i\xi'}{\log T^\gamma} \right|^{\delta-1} \\
&\quad \times \left| \tilde{F}\left(1 + \frac{1+i\xi}{\log T^\gamma}\right) \tilde{F}\left(1 + \frac{1+i\xi'}{\log T^\gamma}\right) \tilde{F}\left(1 + \frac{2+i(\xi+\xi')}{\log T^\gamma}\right)^{-1} \theta(\xi)\theta(\xi') \right| d\xi d\xi' \\
&\ll (\log T)^{\delta-1} \tilde{F}(1) \int_I \int_I (1+|\xi|)^{-2} (1+|\xi'|)^{-2} d\xi d\xi' \\
&\ll (\log T)^{\delta-1} \prod_{\substack{p \in \mathcal{P}_2 \\ p < w(T)}} \left(1 - \frac{1}{p}\right)^{-1}.
\end{aligned}$$

This concludes the proof of (8.13) and so completes the proof of the lemma.  $\square$

**8.3. Conclusion.** Let  $1 \leq i \leq r$ . We are finally in a position to reveal the majorant for the representation function  $R'_i : \{-T, \dots, T\} \rightarrow \mathbb{R}_{\geq 0}$  in (8.1), where  $\mathcal{S}_{C_1, T}$  is the exceptional set from Definition 7.5.

Let  $A \in \mathcal{A}_i$  and let  $W_0 = \prod_{p|D_{K_i}} p^{\alpha(p)}$ . Let  $A' = \gcd(A, W)$  and let  $A_0 = \gcd(A, W_0)$ . Then

$$\gcd\left(\frac{A'}{A_0}, \frac{Wt+A}{A'}\right) = 1,$$

for any  $t \in \mathbb{Z}$ . Hence it follows from Lemma 8.4

$$R_i(Wt+A) \ll \frac{\varrho_i(W_0, A)}{W_0^{n_i-1}} r_{K_i}\left(\frac{A'}{A_0}\right) r_{K_i}\left(\frac{Wt+A}{A'}\right), \quad (8.14)$$

when  $Wt+A > 0$ . Put  $m = Wt+A$  and assume that  $m \notin \mathcal{S}_{C_1, T}$ . Then  $A' = m'$ , in the notation of (8.5). Combining the majorants (8.8) and (8.10) according to (8.7), we obtain

$$\begin{aligned}
r_{K_i}\left(\frac{m}{m'}\right) &\leq r_{i, \text{res}}\left(\frac{m}{m'}\right) \sum_{\substack{q \in \langle \overline{\mathcal{P}}_2^{(i)} \rangle \\ v_p(q) \neq 1 \ \forall p}} \mathbf{1}_{q|m} \tau(q)^{n_i} \mathbf{1}_{\langle \mathcal{P}_0^{(i)} \cup \mathcal{P}_1^{(i)} \rangle} \left(\frac{m}{qm'}\right) \\
&\ll \nu_i^{(T)}(m) \sum_{\substack{q \in \langle \overline{\mathcal{P}}_2^{(i)} \rangle \\ v_p(q) \neq 1 \ \forall p}} \mathbf{1}_{q|m} \tau(q)^{n_i} \chi\left(\frac{\log q}{\log T^\gamma}\right) \nu_{i, \text{sieve}}^{(T)}\left(\frac{m}{q}\right) \\
&= \nu_i^{(T)}(m) \nu'_{i, \text{sieve}}(m).
\end{aligned} \quad (8.15)$$

Here we have noted that  $\nu_{i, \text{sieve}}^{(T)}(m/(qm')) = \nu_{i, \text{sieve}}^{(T)}(m/q)$  and truncated the  $q$  summation using  $\chi$ . To see that the latter is valid, suppose that  $q \mid m$  with  $q > T^{\gamma/2}$  and recall that  $\chi(x) = 1$  for  $x \in [-1/2, 1/2]$ . If there is a prime divisor  $p > T^{1/(\log \log T)^3}$  of  $q$ , then there exists a divisor  $p^2 \mid m$  with  $p^2 > (\log T)^{C_1}$ , since  $q$  is square-full, which implies that  $m$  is rough in the sense of part (1) of Lemma 7.4. If, on the other hand,  $p \leq T^{1/(\log \log T)^3}$  for

every  $p \mid q$ , then  $m$  is smooth in the sense of part (2) of the lemma. Neither case can occur since  $m \notin \mathcal{S}_{C_1, T}$ .

Our final task is to check condition (ii) from the start of Section 7, which states that the mean value of our majorant should agree with the mean value of  $R'_i$ , with respect to  $T$ .

**Lemma 8.9.** *Let  $1 \leq i \leq r$  and define  $\mathcal{A}_i$  as in (6.2). Suppose  $1 \leq A < W$  with  $A \in \mathcal{A}_i$  and write  $A' = \gcd(A, W)$ . Then*

$$\frac{\varrho_i(W_0, A)}{W_0^{n_i-1}} r_{K_i} \left( \frac{A'}{A_0} \right) \mathbb{E}_{m \leq (T-A)/W} \nu_i^{(T)}(Wm + A) \nu'_{i, \text{sieve}}(Wm + A) \asymp \frac{\varrho_i(W, A; M)}{W^{n_i-1}},$$

provided the parameter  $\gamma$  appearing in  $\nu_i^{(T)}$  and  $\nu'_{i, \text{sieve}}(T)$  is sufficiently small.

*Proof.* To begin with, recall that  $\varrho_i(W, A; M) > 0$  and that  $v_p(A) < v_p(M) \ll 1$  for all  $p \mid M$  by our assumption (6.2). Thus, for  $p \mid M$ , the first part of Lemma 4.2 yields

$$0 < \frac{\varrho_i(p^{\alpha(p)}, A; p^{v_p(M)})}{p^{\alpha(p)(n_i-1)}} \leq \frac{\varrho_i(p^{\alpha(p)}, A)}{p^{\alpha(p)(n_i-1)}} \ll 1.$$

Since the above is positive, we may use Lemma 3.4 to deduce a matching lower bound as in (8.4). Thus, the multiplicativity of  $\varrho_i$  implies that

$$\frac{\varrho_i(W, A)}{W^{n_i-1}} \ll \frac{\varrho_i(W, A; M)}{W^{n_i-1}} \leq \frac{\varrho_i(W, A)}{W^{n_i-1}}. \quad (8.16)$$

Next we note that  $A'/A_0 = \gcd(W/W_0, A)$ . Hence the second part of Lemma 4.2 yields

$$\frac{\varrho_i(W, A)}{W^{n_i-1}} = \frac{\varrho_i(W_0, A)}{W_0^{n_i-1}} r_{K_i} \left( \frac{A'}{A_0} \right) \prod_{\substack{p < w(T) \\ p \nmid D_{K_i}}} \left( 1 - \frac{1}{p} \right)^{-1} \prod_{\substack{\mathfrak{p} \mid p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left( 1 - \frac{1}{N \mathfrak{p}} \right). \quad (8.17)$$

This reduces our task to establishing, for sufficiently small  $\gamma$ , the estimate

$$\mathbb{E}_{m \leq (T-A)/W} \nu_i^{(T)}(Wm + A) \nu'_{i, \text{sieve}}(Wm + A) \asymp \prod_{p < w(T)} \left( 1 - \frac{1}{p} \right)^{-1} \prod_{\substack{\mathfrak{p} \mid p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left( 1 - \frac{1}{N \mathfrak{p}} \right).$$

We temporarily set

$$S(T') = \mathbb{E}_{m \leq T'/W} \nu_i^{(T)}(Wm + A) \nu'_{i, \text{sieve}}(Wm + A),$$

for any  $T' \leq T$ . The two factors of the majorant are truncated divisor sums. According to the discussion following (8.8), the first function  $\nu_i^{(T)}$  is constructed from divisors in  $\langle \overline{\mathcal{P}}_1^{(i)} \rangle$  provided  $T$  is sufficiently large, whereas the function  $\nu'_{i, \text{sieve}}(T)$  is constructed from divisors belonging to  $\langle \overline{\mathcal{P}}_2^{(i)} \rangle$ . In particular the divisors used in the construction of the former are all coprime to the divisors appearing in the latter. We therefore deduce (cf. [27, p. 262]) that

$$\begin{aligned} S(T') &= \mathbb{E}_{m \leq T'/W} \nu_i^{(T)}(Wm + A) \\ &\quad \times \mathbb{E}_{m \leq T'/W} \nu'_{i, \text{sieve}}(Wm + A) + O(T^{O(\gamma)} W/T'), \end{aligned}$$

which provides an asymptotic formula whenever  $T^{O(\gamma)}W = o(T')$ . Combining Lemmas 8.7 and 8.8, this allows us to deduce the upper bound

$$S(T - A) \ll \prod_{p < w(T)} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{\mathfrak{p}|p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left(1 - \frac{1}{N\mathfrak{p}}\right)$$

if  $\gamma$  is sufficiently small. To obtain the lower bound for  $S(T - A)$ , we combine (8.14), (8.15), (8.3) to get

$$\begin{aligned} \frac{\varrho_i(W_0, A)}{W_0^{n_i-1}} r_{K_i} \left( \frac{A'}{A_0} \right) S(T - A) &\gg \mathbb{E}_{m \leq (T-A)/W} R'_i(Wm + A) \\ &\gg \frac{\varrho_i(W, A; M)}{W^{n_i-1}}. \end{aligned}$$

But then it follows that

$$S(T - A) \gg \prod_{p < w(T)} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{\mathfrak{p}|p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left(1 - \frac{1}{N\mathfrak{p}}\right),$$

by (8.16) and (8.17).  $\square$

For every  $i \in \{1, \dots, r\}$  let  $A_i$  be such that  $1 \leq A_i < W$  and  $A_i \pmod{W} \in \mathcal{A}_i$ . The proof of Lemma 8.9 shows that there is a function

$$\varphi_i(T; A_i) \asymp \prod_{p < w(T)} \left(1 - \frac{1}{p}\right) \prod_{\substack{\mathfrak{p}|p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left(1 - \frac{1}{N\mathfrak{p}}\right)^{-1} \quad (8.18)$$

such that

$$\varphi_i(T; A_i) \mathbb{E}_{m < T/W} \nu_i^{(T)}(Wm + A_i) \nu'_{i, \text{sieve}}(Wm + A_i) = 1. \quad (8.19)$$

We define the joint normalised majorant function

$$\varpi_{A_1, \dots, A_r}^{(T)}(m) = \frac{1}{r} \sum_{i=1}^r \varphi_i(T; A_i) \nu_i^{(T)}(Wm + A_i) \nu'_{i, \text{sieve}}(Wm + A_i). \quad (8.20)$$

We will often write

$$\varpi^{(T)}(m) = \varpi_{A_1, \dots, A_r}^{(T)}(m),$$

for short. It satisfies  $\mathbb{E}_{m < T/W} \varpi^{(T)}(m) = 1$ . Moreover,  $\varpi^{(T)}$  simultaneously majorises the *normalised* counting functions

$$\left( \frac{\varrho_i(W, A_i; M)}{W^{n_i-1}} \right)^{-1} R'_i(Wm + A_i),$$

for  $1 \leq i \leq r$  and  $R'_i$  as in (8.1), in the sense of (i) from the start of Section 7.

## 9. THE MAJORANT IS PSEUDORANDOM

Let  $A_i \in \mathcal{A}_i$  for  $1 \leq i \leq r$ , in the notation of (6.2), and recall the definition (8.20) of  $\varpi^{(T)} = \varpi_{A_1, \dots, A_r}^{(T)}$ . Given  $D > 1$ , our aim in this section is to show that the family  $(\varpi^{(T)})_{T \in \mathbb{Z}_{>0}}$  gives rise to a family of  $D$ -pseudorandom majorants, in the sense of [18, §6] with  $m_0 = d_0 = L_0 = D$ , provided that the parameter  $\gamma$  appearing in the truncations is sufficiently small. In our setting it suffices to consider  $D \ll_L 1$ , where  $L$  is as in (5.3).

For each  $T$  let  $\tilde{T}$  be a prime number such that  $T/W < \tilde{T} \ll_L T/W$ . Choosing  $\tilde{T}$  sufficiently large in terms of  $L$  allows us to pass from counting problems within the set of integers  $\{1, \dots, [T/W]\}$  to counting problems in the group  $\mathbb{Z}/\tilde{T}\mathbb{Z}$ , without creating new solutions due to the wrap-around effect. The majorants are extended to  $\mathbb{Z}/\tilde{T}\mathbb{Z}$  by defining  $\varpi'^{(T)} : \mathbb{Z}/\tilde{T}\mathbb{Z} \rightarrow \mathbb{R}_{>0}$  via

$$\varpi'^{(T)}(m) = \begin{cases} (1 + \varpi^{(T)}(m))/2, & \text{if } m \leq T/W, \\ 1, & \text{if } T/W < m \leq \tilde{T}. \end{cases}$$

By [18, App. D] it suffices to prove the following two propositions in order to show that  $(\varpi'^{(T)})_{T \in \mathbb{Z}_{>0}}$  is a family of  $D$ -pseudorandom majorants. As indicated above, we will apply them with  $D \ll_L 1$ .

**Proposition 9.1** (*D*-Linear forms estimate). *Let  $T \in \mathbb{Z}_{>0}$ ,  $T' = \lfloor \frac{T}{W} \rfloor$ , and let  $D > 1$ . Suppose that  $1 \leq r', s' \leq D$  and let  $\mathbf{h} = (h_1, \dots, h_{r'}) : \mathbb{Z}^{s'} \rightarrow \mathbb{Z}$  be a system of linear polynomials whose non-constant parts are pairwise non-proportional. Suppose that coefficients of each  $h_i$ , other than possibly the constant terms, are bounded in absolute value by  $D$ , while  $h_i(0) = O_D(T)$ . Suppose  $\mathfrak{K} \subset [-1, 1]^{s'}$  is a convex body such that  $\mathbf{h}(T'\mathfrak{K}) \subset [1, T']^{r'}$  and  $\text{vol}(\mathfrak{K}) \gg 1$ . Then we have*

$$\frac{1}{\text{vol}(T'\mathfrak{K})} \sum_{\mathbf{m} \in \mathbb{Z}^{s'} \cap T'\mathfrak{K}} \prod_{j=1}^{r'} \varpi_{A_1, \dots, A_r}^{(T)}(h_j(\mathbf{m})) = 1 + o_D(1), \quad (9.1)$$

provided  $\gamma$  is small enough.

**Proposition 9.2** (*D*-correlation estimate). *Let  $T \in \mathbb{Z}_{>0}$ ,  $T' = \lfloor \frac{T}{W} \rfloor$  and let  $D > 1$ . Then there exists a function  $\sigma : \{-T', \dots, T'\} \rightarrow \mathbb{R}_{\geq 0}$  with bounded moments*

$$\mathbb{E}_{|m| \leq T'} \sigma^q(m) \ll_{D,q} 1,$$

such that for every discrete interval  $I \subset \{1, \dots, T'\}$ , every  $1 \leq d \leq D$ , every  $(i_1, \dots, i_d) \in \{1, \dots, r\}^d$  and every choice of (not necessarily distinct)  $a_1, \dots, a_d \in \{1, \dots, T'\}$ , we have

$$\begin{aligned} \sum_{m \in I} \prod_{j=1}^d \varphi_{i_j}(T; A_{i_j}) \nu_{i_j}^{(T)}(W(m + a_j) + A_{i_j}) \nu'_{i_j, \text{sieve}}^{(T)}(W(m + a_j) + A_{i_j}) \\ \leq T' \sum_{1 \leq j < j' \leq d} \sigma(a_j - a_{j'}), \end{aligned}$$

provided  $\gamma$  is small enough.

In proving Propositions 9.1 and 9.2, we will allow all of our implied constants to depend on the parameter  $D$ . We begin with the proof of the former. Unravelling definitions, we see that (9.1) is implied by the estimate

$$\begin{aligned} & \frac{1}{\text{vol}(T' \mathfrak{K})} \sum_{\mathbf{m} \in \mathbb{Z}^{s'} \cap T' \mathfrak{K}} \prod_{j=1}^{r'} \nu_{i_j}^{(T)}(Wh_j(\mathbf{m}) + A_{i_j}) \nu'_{i_j, \text{sieve}}(Wh_j(\mathbf{m}) + A_{i_j}) \\ &= (1 + o(1)) \prod_{j=1}^{r'} \varphi_{i_j}(T; A_{i_j})^{-1}, \end{aligned} \quad (9.2)$$

for every collection of indices  $1 \leq i_1, \dots, i_{r'} \leq r$ . Here we have  $\varphi_{i_j}(T; A_{i_j})^{-1} \asymp \Pi_{i_j}$ , by (8.18), where

$$\Pi_i = \prod_{p < w(T)} \left(1 - \frac{1}{p}\right)^{-1} \prod_{\substack{\mathfrak{p} | p \\ \mathfrak{p} \subset \mathfrak{o}_{K_i}}} \left(1 - \frac{1}{N \mathfrak{p}}\right). \quad (9.3)$$

The strategy to proving (9.2) is the same as in [27, §9], which is related to that of [18, App. D]. Inserting all definitions and writing  $g_i = \mu * r_{i, \text{res}}$ , we have

$$\begin{aligned} & \nu_i^{(T)}(Wm + A) \nu'_{i, \text{sieve}}(Wm + A) \\ &= \sum_{\kappa=4/\gamma}^{[(\log \log T)^3]} \sum_{\lambda=\lceil \log_2 \kappa - 2 \rceil}^{[\log_2((\log \log T)^3)]} \sum_{u \in U(\lambda, \kappa)} 2^{n_i \kappa} \mathbf{1}_{u|Wm+A} r_{i, \text{res}}(u) \\ & \times \sum_{\substack{d \in \langle \overline{\mathcal{P}}_1^{(i)} \rangle \\ \gcd(d, u)=1}} \mathbf{1}_{d|Wm+A} g_i(d) \chi\left(\frac{\log d}{\log T^\gamma}\right) \\ & \times \sum_{\substack{q \in \langle \overline{\mathcal{P}}_2^{(i)} \rangle \\ v_p(q) \neq 1 \ \forall p}} \mathbf{1}_{q|Wm+A} \tau(q)^{n_i} \chi\left(\frac{\log q}{\log T^\gamma}\right) \left( \sum_{\substack{e \in \langle \overline{\mathcal{P}}_2^{(i)} \rangle \\ qe|Wm+A}} \mu(e) \chi\left(\frac{\log e}{\log T^\gamma}\right) \right)^2. \end{aligned}$$

Here the restriction to  $d \in \langle \overline{\mathcal{P}}_1^{(i)} \rangle$  arises from Remark 7.8 and the fact that  $g_i(d) = 0$  when  $d$  has a prime factor  $p < w(T)$ . Noting that  $\gcd(qe, d) = 1$ , the right hand side is seen to be

$$\begin{aligned} & \sum_{\kappa=4/\gamma}^{[(\log \log T)^3]} \sum_{\lambda=\lceil \log_2 \kappa - 2 \rceil}^{[\log_2((\log \log T)^3)]} \sum_{u \in U(\lambda, \kappa)} \sum_{\substack{d \in \langle \overline{\mathcal{P}}_1^{(i)} \rangle \\ \gcd(d, u)=1}} \sum_{\substack{q \in \langle \overline{\mathcal{P}}_2^{(i)} \rangle \\ v_p(q) \neq 1 \ \forall p}} \\ & \times \sum_{e, e' \in \langle \overline{\mathcal{P}}_2^{(i)} \rangle} 2^{n_i \kappa} r_{i, \text{res}}(u) \tau(q)^{n_i} \mu(e) \mu(e') g_i(d) \mathbf{1}_{\Delta|Wm+A} \prod_{x \in \{d, e, e', q\}} \chi\left(\frac{\log x}{\log T^\gamma}\right), \end{aligned}$$

where  $\Delta = \text{lcm}(u, d, qe, qe')$ . Together, Remark 7.7 and the compact support of  $\chi$  ensure that all divisors  $d, e, e', q, u$  are bounded by  $T^\gamma$ . For each  $1 \leq j \leq r'$  we define the linear

polynomial

$$h'_j(\mathbf{m}) = Wh_j(\mathbf{m}) + A_{i_j}.$$

We may assume that  $T$  is sufficiently large in terms of  $D$  to ensure that the non-constant parts of the polynomials  $h_1, \dots, h_{r'}$  are pairwise non-proportional modulo any prime  $p > w(T)$ . The same then holds for the polynomials  $h'_1, \dots, h'_{r'}$ .

Let  $\Delta_j = \text{lcm}(u_j, d_j, q_j e_j, q_j e'_j)$ , for  $1 \leq j \leq r'$ . We are interested in estimating the cardinality

$$\begin{aligned} & \#\{\mathbf{m} \in \mathbb{Z}^{s'} \cap T' \mathfrak{K} : \Delta_j \mid h'_j(\mathbf{m})\} \\ &= \sum_{\substack{\mathbf{s} \pmod{\Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'}} \\ \Delta_j \mid h'_j(\mathbf{m})}} \#\{\mathbf{m} \in \mathbb{Z}^{s'} \cap T' \mathfrak{K} : \mathbf{m} \equiv \mathbf{s} \pmod{\Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'}}\}, \end{aligned}$$

where  $\Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} = \text{lcm}(\Delta_1, \dots, \Delta_{r'}) \leq T^{O(\gamma)}$ . Extending the notion of local divisor densities multiplicatively from (5.5), with  $\mathcal{W}_m = (\mathbb{Z}/p^m \mathbb{Z})^{s'}$  and the set of polynomials  $\mathbf{h}'$ , the outer sum has cardinality  $\alpha_{\mathbf{h}'}(\Delta_1, \dots, \Delta_{r'}) \Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'}^{s'}$ . The inner cardinality is equal to

$$\# \left( (\Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \mathbb{Z}^{s'} + \mathbf{s}) \cap T' \mathfrak{K} \right) = \# \left( \mathbb{Z}^{s'} \cap (\Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'}^{-1} T' \mathfrak{K} + \Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'}^{-1} \mathbf{s}) \right).$$

We may therefore apply Lemma 3.1 with  $\mathcal{B} = \mathfrak{K}$  and  $T = \Delta_{\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'}^{-1} T'$  to each of the above cardinalities. This leads to the conclusion that

$$\begin{aligned} & \frac{1}{\text{vol}(T' \mathfrak{K})} \sum_{\mathbf{m} \in \mathbb{Z}^{s'} \cap T' \mathfrak{K}} \prod_{j=1}^{r'} \nu_{i_j}^{(T)}(h'_j(\mathbf{m})) \nu'_{i_j, \text{sieve}}^{(T)}(h'_j(\mathbf{m})) \\ &= \sum_{\boldsymbol{\kappa}} \sum_{\boldsymbol{\lambda}} \sum_{\substack{\mathbf{u} \\ u_j \in U(\lambda_j, \kappa_j)}} \sum_{\substack{\mathbf{d} \\ \gcd(d_j, u_j)=1}} \sum_{\mathbf{q}} \sum_{\mathbf{e}, \mathbf{e}'} \left( \alpha_{\mathbf{h}'}(\Delta_1, \dots, \Delta_{r'}) + O\left(\frac{T'^{-1+O(\gamma)}}{\text{vol}(\mathfrak{K})}\right) \right) \quad (9.4) \\ & \quad \times \prod_{j=1}^{r'} 2^{\kappa_j n_{i_j}} \tau(q_j)^{n_{i_j}} \mu(e_j) \mu(e'_j) g_{i_j}(d_j) \prod_{x \in \{d_j, e_j, e'_j, q_j\}} \chi\left(\frac{\log x}{\log T^\gamma}\right), \end{aligned}$$

where  $\boldsymbol{\kappa}, \boldsymbol{\lambda}, \mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \in \mathbb{Z}^{r'}$  are assumed to satisfy the correct multiplicative restrictions component-wise. Thus, for example, the sum over  $\mathbf{d}$  is restricted to

$$\{\mathbf{d} : d_j \in \langle \overline{\mathcal{P}}_1^{(i_j)} \rangle, 1 \leq j \leq r'\}.$$

Similarly, those over  $\mathbf{q}, \mathbf{e}$  and  $\mathbf{e}'$  are restricted to

$$\{\mathbf{x} : x_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle, 1 \leq j \leq r'\}.$$

We assume, furthermore, that all coordinates of  $\mathbf{q}$  satisfy  $v_p(q_j) \neq 1$  for all primes  $p$ .

We begin by examining the error term in (9.4). As mentioned above, each of the sums over  $u_j, d_j, e_j, e'_j$  and  $q_j$  have at most  $T^\gamma$  terms. Together with the trivial bounds  $\tau(q) \leq T^\gamma$

for  $q \leq T^\gamma$  and  $2^{\kappa_j} \leq 2^{(\log \log T)^3} \ll T^{o(1)}$ , this implies that the error term makes a total contribution of

$$O\left(\frac{T'^{-1+O(\gamma)}}{\text{vol } \mathfrak{K}}\right) = o(1),$$

by our assumptions on  $\mathfrak{K}$ .

The main term will now be analysed in much the same way as in [26, §6] and [27, §9]. Our majorant very closely resembles that from [27], the latter in fact being a special case of it. The analysis of (9.4) is therefore only a minor adaptation of what is established in [27, §9]. Given the length of the argument we include an overview here as guidance, and only include the details of the more complicated proofs where it may not be immediately clear that the corresponding argument from [27, §9] still applies.

Any prime  $p \mid \Delta$  satisfies  $p > w(T)$ . Hence  $\alpha_{\mathbf{h}'}(\Delta_1, \dots, \Delta_{r'})$  will be determined using the first three alternatives from (5.6). In particular,

$$\alpha_{\mathbf{h}'}(\Delta_1, \dots, \Delta_{r'}) = \prod_{j=1}^{r'} \frac{1}{\Delta_j}$$

whenever  $\Delta_1, \dots, \Delta_{r'}$  are pairwise coprime. Put  $\tilde{\Delta}_j = d_j q_j \text{lcm}(e_j, e'_j)$ , for  $1 \leq j \leq r'$ . The first step is to show that we may replace  $\alpha_{\mathbf{h}'}(\Delta_1, \dots, \Delta_{r'})$  in the main term by

$$\frac{\alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'})}{u_1 \dots u_{r'}},$$

at the expense of an overall error term  $o(1)$ . To prove this, it suffices to show that we may restrict the summation to vectors  $(\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}')$  for which

$$\gcd(u_j, \tilde{\Delta}_j) = 1 \quad \text{and} \quad \gcd(u_i, u_j \tilde{\Delta}_j) = 1$$

for all  $j$  and all  $i \neq j$ . Since  $\gcd(d_j, u_j) = 1$ , it follows from Remark 7.8 that the first condition is always satisfied. Furthermore, the set of all vectors  $(\mathbf{u}, \mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}')$  failing the second condition makes a negligible contribution. The proof follows (cf. the proof of [27, Claim 2]) by the Cauchy–Schwarz inequality from a second-moment estimate together with a lower bound on the prime divisors of any  $u_j$ .

We note that

$$\sum_{\kappa} \sum_{\lambda} \sum_{u_j \in U(\lambda_j, \kappa_j)} \prod_{j=1}^{r'} \frac{2^{\kappa_j n_{i_j}} r_{j, \text{res}}(u_j)}{u_j} < \infty, \quad (9.5)$$

The absolute convergence of this sum follows from the proof of Proposition 7.6.

The next step is to replace  $\chi$  by a multiplicative function using (8.12). For  $1 \leq j \leq r'$  and  $1 \leq k \leq 4$  we write

$$z_{j,k} = \frac{1 + i\xi_{j,k}}{\log T^\gamma}.$$

Likewise we set  $d\xi = \prod_{j,k} d\xi_{j,k}$  and

$$J_j = \mu(e_j) \mu(e'_j) e_j^{-z_{j,1}} e_j'^{-z_{j,2}} g_{i_j}(d_j) d_j^{-z_{j,3}} \tau(q_j)^{n_{i_j}} q_j^{-z_{j,4}}, \quad (9.6)$$



for  $1 \leq j \leq r'$ . With this notation the new main term is equal to

$$\sum_{\kappa} \sum_{\lambda} \sum_{\mathbf{u}} \prod_{j=1}^{r'} \frac{2^{\kappa_j n_{i_j}} r_{j,\text{res}}(u_j)}{u_j} \left( J(\mathbf{u}) + O_E \left( \frac{1}{(\log T)^E} \sum_{(\mathbf{d}, u_j)=1} \sum_{\mathbf{q}} \sum_{\mathbf{e}, \mathbf{e}'} \prod_{j=1}^{r'} \frac{H^{\Omega(d_j q_j)} \alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'})}{(e_j e'_j d_j q_j)^{1/\log T^\gamma}} \right) \right)$$

for any  $E > 0$ , where  $H = \max_{1 \leq i \leq r'} 2^{n_i}$  and

$$J(\mathbf{u}) = \sum_{(\mathbf{d}, u_j)=1} \sum_{\mathbf{q}} \sum_{\mathbf{e}, \mathbf{e}'} \alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'}) \int_I \dots \int_I \left( \prod_{j=1}^{r'} J_j \prod_{k=1}^4 \vartheta(\xi_{j,k}) \right) d\xi.$$

The error terms that appear in the next step will again depend on how small the prime factors of the relevant numbers can be. This time these are the coordinates of  $\mathbf{d}, \mathbf{q}, \mathbf{e}$  and  $\mathbf{e}'$  instead of  $\mathbf{u}$  and we can only assume that the primes are larger than  $w(T)$ , which is much smaller than the lower bound on prime factors of the  $u_i$ . For this reason it was essential to treat the  $u_j$  separately first, in order to make use of the convergence of the sums over  $\kappa, \lambda$  and  $U(\lambda_j, \kappa_j)$  when showing that the new error term is negligible.

The next step is to show that we may swap the product over  $j$  with all the sums. That is, we replace  $\alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'})$  by  $(\tilde{\Delta}_1 \dots \tilde{\Delta}_{r'})^{-1}$ , while only introducing a small error. We will show that

$$\begin{aligned} J(\mathbf{u}) + o \left( \prod_{j=1}^{r'} \Pi_{i_j} \right) &= \int_I \dots \int_I \left( \prod_{j=1}^{r'} \sum_{\substack{d_j \\ (d_j, u_j)=1}} \sum_{\substack{q_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle \\ v_p(q_j) \neq 1 \ \forall p}} \sum_{\substack{e_j, e'_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle}} \frac{J_j}{\tilde{\Delta}_j} \prod_{k=1}^4 \vartheta(\xi_{j,k}) \right) d\xi \\ &= \prod_{j=1}^{r'} \sum_{\substack{d_j \\ (d_j, u_j)=1}} \sum_{\substack{q_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle \\ v_p(q_j) \neq 1 \ \forall p}} \sum_{\substack{e_j, e'_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle}} \int_{I^4} \left( \frac{J_j}{\tilde{\Delta}_j} \prod_{k=1}^4 \vartheta(\xi_{j,k}) \right) d\xi_{j,1} \dots d\xi_{j,4}, \end{aligned} \tag{9.7}$$

where  $\Pi_i$  is given by (9.3) for  $1 \leq i \leq r$ . Before establishing this estimate, we remark that the final main term is now a product of  $r'$  factors that are independent of each other and independent of the system  $\mathbf{h}$  of linear polynomials that we started with. In particular, we may consider this estimate in the special case where  $s' = r' = 1$  and where  $h_1(m) = m$ . Reinstating the sums over  $\kappa, \lambda$  and  $\mathbf{u}$ , this relates the  $j$ th factor of the above product to the average value of the majorant function. By (8.19) and (9.5), we therefore deduce that

$$\sum_{\kappa} \sum_{\lambda} \sum_{\mathbf{u}} \prod_{j=1}^{r'} \frac{2^{\kappa_j n_{i_j}} r_{j,\text{res}}(u_j)}{u_j} \left( J(\mathbf{u}) + o \left( \prod_{j=1}^{r'} \Pi_{i_j} \right) \right) = (1 + o(1)) \prod_{j=1}^{r'} \varphi_{i_j}(T; A_{i_j})^{-1}.$$

This completes the proof of (9.2), and hence the proof of Proposition 9.1, subject to the verification of (9.7). Our proof of (9.7) will be undertaken in two steps, as recorded in the following two results. We fix values of  $\kappa$  and  $\lambda$  for now.

**Lemma 9.3.** *For each  $1 \leq j \leq r'$  and each  $u_j \in U(\lambda_j, \kappa_j)$ , we have*

$$\int_{I^4} \left| \sum_{\substack{d_j \\ \gcd(d_j, u_j)=1}} \sum_{\substack{q_j \in \langle \overline{\mathcal{D}}_2^{(i_j)} \rangle \\ v_p(q_j) \neq 1 \ \forall p}} \sum_{e_j, e'_j \in \langle \overline{\mathcal{D}}_2^{(i_j)} \rangle} \frac{J_j}{\tilde{\Delta}_j} \prod_{k=1}^4 \vartheta(\xi_{j,k}) \right| d\xi_{j,1} \dots d\xi_{j,4} \ll \Pi_{i_j},$$

where  $\Pi_{i_j}$  is given by (9.3), and where the implied constant is independent of  $\lambda_j, \kappa_j, u_j$ .

This lemma corresponds to [27, Claim 5]. We take the opportunity to provide a full proof here, since the extra factors  $g_{i_j}(d_j)\tau(q_j)^{n_{i_j}}$  implicit in  $J_j$  make the analysis slightly more delicate. Moreover, while the proof of [27, Claim 5] is correct, it requires an application of Lemma 3.2, which is not present in [27].

*Proof of Lemma 9.3.* The first step is to express the integrand, which we denote by  $K = K(\xi_{j,1}, \dots, \xi_{j,4})$ , as an Euler product. The fact that  $\Re(z_{j,k}) = (\log T^\gamma)^{-1} > 0$  will allow us to restrict to the square-free part. Recall that

$$\frac{J_j}{\tilde{\Delta}_j} = \frac{\mu(e_j)\mu(e'_j)e_j^{-z_{j,1}}e_j'^{-z_{j,2}}g_{i_j}(d_j)d_j^{-z_{j,3}}\tau(q_j)^{n_{i_j}}q_j^{-z_{j,4}}}{d_jq_j \operatorname{lcm}(e_j, e'_j)}$$

and put

$$\tilde{L} = \prod_{k=1}^4 (1 + |\xi_{j,k}|).$$

By (8.11) we have

$$\begin{aligned} K &\ll_E \tilde{L}^{-E} \left| \prod_{p \in \overline{\mathcal{D}}_2^{(i_j)}} (1 - p^{-1-z_{j,1}} - p^{-1-z_{j,2}} + p^{-1-z_{j,1}-z_{j,2}} + O(p^{-2})) \right| \\ &\quad \times \left| \prod_{\substack{p \in \overline{\mathcal{D}}_1^{(i_j)} \\ p \nmid u}} (1 + (r_{K_{i_j}}(p) - 1)p^{-1-z_{j,3}} + O(p^{-2})) \right|, \end{aligned}$$

for any  $E > 3$ . Since for sufficiently large primes  $p$  each of the factors of these products can be analysed via the logarithmic series, we deduce that

$$\begin{aligned} K &\ll \tilde{L}^{-E} \left| \prod_{p \in \overline{\mathcal{D}}_2^{(i_j)}} (1 - p^{-1-z_{j,1}})(1 - p^{-1-z_{j,2}})(1 + p^{-1-z_{j,1}-z_{j,2}}) \right| \\ &\quad \times \left| \prod_{p \in \overline{\mathcal{D}}_1^{(i_j)}} (1 - p^{-1-z_{j,3}})(1 + r_{K_{i_j}}(p)p^{-1-z_{j,3}}) \right| \prod_{p|u} \left( 1 + \frac{r_{K_{i_j}}(p) + 1}{p} \right). \end{aligned}$$

Since any prime divisor of  $u$  comes from an interval of the form  $[y, y^2]$ , with  $y = T^{1/(2^{\lambda+1})}$ , the final product over  $p|u$  is easily seen to be  $O_{n_{i_j}}(1)$ , and can be ignored.

We will now proceed as in Section 8.2. Let  $1 \leq i \leq r$ . Recall that there are functions  $G_1, G_2$ , which are non-zero and holomorphic on  $\Re(s) \geq 1$ , such that

$$F_{i,1}(s) = \prod_{p \in \mathcal{P}_1^{(i)}} \left(1 - \frac{1}{p^s}\right)^{-1} = \zeta^{\delta_i}(s) G_1(s),$$

$$F_{i,2}(s) = \prod_{p \in \mathcal{P}_2^{(i)}} \left(1 - \frac{1}{p^s}\right)^{-1} = \zeta^{1-\delta_i}(s) G_2(s).$$

Hence, when  $\Re s > 0$  and  $|s| \ll 1$ , then

$$|F_{i,1}(1+s)| \asymp |s|^{-\delta_i}, \quad |F_{i,2}(1+s)| \asymp |s|^{-1+\delta_i}.$$

Likewise,  $|F_{i,1}^{-1}(1+s)| \asymp |s|^{\delta_i}$  and  $|F_{i,2}^{-1}(1+s)| \asymp |s|^{1-\delta_i}$ . In order to employ these asymptotic orders to bound the integral above, we apply Lemma 3.2 to deduce that there is an absolute positive constant  $C$  such that each of the three Euler products  $E(s)$ , given by

$$H_{i,1}(s) = \prod_{\substack{p \in \mathcal{P}_1^{(i)} \\ C < p < w(T)}} \left(1 - \frac{1}{p^s}\right)^{-1}, \quad H_{i,2}(s) = \prod_{\substack{p \in \mathcal{P}_2^{(i)} \\ C < p < w(T)}} \left(1 - \frac{1}{p^s}\right)^{-1},$$

and

$$H_{K_i}(s) = \prod_{C < p < w(T)} \left(1 + \frac{r_{K_i}(p)}{p^s}\right)^{-1},$$

satisfies  $|E(1+s)| \asymp E(1)$  and (3.1) when  $|s| \ll (\log T^\gamma)^{-1/2}$ . Recall the definition (9.3) of  $\Pi_{i_j}$ . We may conclude that

$$\begin{aligned} K &\ll \tilde{L}^{-E} |z_{j,1}|^{1-\delta_{i_j}} |z_{j,2}|^{1-\delta_{i_j}} |z_{j,1} + z_{j,2}|^{\delta_{i_j}-1} |z_{j,3}|^{\delta_{i_j}} |z_{j,3}|^{-1} \\ &\quad \times H_{i_j,2}(1) H_{i_j,1}(1) \prod_{p < w(T)} \left(1 - \frac{r_{K_i}(p)}{p}\right) \\ &\ll \tilde{L}^{-E} \left(\frac{1 + |\xi_{j,1}|}{\log T^\gamma}\right)^{1-\delta_{i_j}} \left(\frac{1 + |\xi_{j,2}|}{\log T^\gamma}\right)^{1-\delta_{i_j}} \left(\frac{1 + |\xi_{j,1} + \xi_{j,2}|}{\log T^\gamma}\right)^{\delta_{i_j}-1} \\ &\quad \times \left(\frac{1 + |\xi_{j,3}|}{\log T^\gamma}\right)^{\delta_{i_j}-1} \Pi_{i_j} \\ &\ll \tilde{L}^{-E/2} \Pi_{i_j}, \end{aligned}$$

since  $0 < \delta_{i_j} \leq 1$ . The lemma now follows since  $\int_{I^4} \tilde{L}^{-E/2} d\xi_{j,1} \dots d\xi_{j,4} = O_E(1)$ .  $\square$

**Lemma 9.4.** *For every  $\mathbf{u} \in U(\kappa, \lambda)$ , we have*

$$\sum_{\substack{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \\ \gcd(d_j, u_j)=1}} \alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'}) J_1 \dots J_{r'} = (1 + o(1)) \prod_{j=1}^{r'} \sum_{\substack{d_j: \\ (d_j, u_j)=1}} \sum_{\substack{q_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle \\ v_p(q_j) \neq 1 \ \forall p}} \sum_{\substack{e_j, e'_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle}} \frac{J_j}{\tilde{\Delta}_j}.$$

Before establishing this result, let us indicate how it suffices to conclude the proof of (9.7). The second equality in (9.7) is obvious, and so only the first requires a proof. Lemma 9.4 implies that the difference of the two integrands is pointwise bounded by

$$o\left(\left|\prod_{j=1}^{r'} \sum_{\substack{d_j: \\ (d_j, u_j)=1}} \sum_{\substack{q_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle \\ v_p(q_j) \neq 1 \ \forall p}} \sum_{\substack{e_j, e'_j \in \langle \overline{\mathcal{P}}_2^{(i_j)} \rangle}} \frac{J_j}{\tilde{\Delta}_j}\right|\right).$$

Lemma 9.3 implies that the integral over this bound equals  $o(\prod_j \Pi_{i_j})$ , which implies the first part of (9.7).

*Proof of Lemma 9.4.* Our argument is identical to that of [27, Claim 3], but we provide more detail here. Throughout this proof we assume, without explicitly mentioning so, that all entries  $d_j$  of any vector  $\mathbf{d}$  satisfy  $\gcd(d_j, u_j) = 1$ . The aim is to study the multiplicative function

$$\eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') = \alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'}) J_1 \dots J_{r'},$$

where  $J_j$  is given by (9.6). We may factorise

$$\eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') = \eta(\tilde{\mathbf{d}}, \tilde{\mathbf{q}}, \tilde{\mathbf{e}}, \tilde{\mathbf{e}}') \eta(\overline{\mathbf{d}}, \overline{\mathbf{q}}, \overline{\mathbf{e}}, \overline{\mathbf{e}}')$$

in such a way that in the first factor the  $r'$  entries  $\tilde{d}_j \tilde{q}_j \tilde{e}_j \tilde{e}'_j$  for  $1 \leq j \leq r'$  of  $\alpha_{\mathbf{h}'}$  are pairwise coprime, while in the second factor, any prime that divides one entry of  $\alpha_{\mathbf{h}'}$  also divides a second entry. The aim is to show that the main contribution from either side of the expression in the statement of the lemma comes from such vectors  $(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}')$  for which the second factor in this decomposition is  $1 = \eta(\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1})$ .

We begin with the left hand side. Let  $k$  be an integer. Then, with the notation  $p^{\mathbf{x}} = (p^{x_1}, \dots, p^{x_{r'}})$  for a prime  $p$  and  $\mathbf{x} \in \mathbb{Z}_{\geq 0}^{r'}$ , we have

$$\sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') = \left( \sum'_{\substack{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \\ \gcd(d_j q_j e_j e'_j, k)=1 \\ 1 \leq j \leq r'}} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \right) \prod_{p|k} \left( \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \in \mathbb{Z}_{\geq 0}^{r'}} \eta(p^{\mathbf{d}}, p^{\mathbf{q}}, p^{\mathbf{e}}, p^{\mathbf{e}'} ) \right),$$

where  $\sum'$  denotes that the sum is restricted to coprime vectors in the above sense; i.e.  $(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') = (\tilde{\mathbf{d}}, \tilde{\mathbf{q}}, \tilde{\mathbf{e}}, \tilde{\mathbf{e}}')$  in the first two sums, and in the third sum only one of the  $r'$  integers  $\max(d_j, q_j, e_j, e'_j)$  may be non-zero. We claim that for  $p > w(T)$  and sufficiently large  $T$  the latter sum satisfies

$$\sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \in \mathbb{Z}_{\geq 0}^{r'}} \eta(p^{\mathbf{d}}, p^{\mathbf{q}}, p^{\mathbf{e}}, p^{\mathbf{e}'} ) = 1 + O(p^{-1}). \quad (9.8)$$

Taking this on trust for a moment, we see that the previous two equations imply

$$\sum'_{\substack{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \\ \gcd(d_j q_j e_j e'_j, k)=1 \\ 1 \leq j \leq r'}} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') = \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \prod_{p|k} (1 + O(p^{-1})).$$

Applying this with  $k = \prod_{j=1}^{r'} \bar{d}_j \bar{q}_j \bar{e}_j \bar{e}'_j$ , we obtain

$$\begin{aligned} & \sum_{\substack{(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \\ \{d_j q_j e_j e'_j : 1 \leq j \leq r'\} \\ \text{not pairwise coprime}}} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \\ &= \left( \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \right) \left( \sum_{\substack{(\bar{\mathbf{d}}, \bar{\mathbf{q}}, \bar{\mathbf{e}}, \bar{\mathbf{e}}') \\ \neq (1, 1, 1, 1)}}^{\dagger} \eta(\bar{\mathbf{d}}, \bar{\mathbf{q}}, \bar{\mathbf{e}}, \bar{\mathbf{e}}') \prod_{p|k} (1 + O(p^{-1})) \right), \end{aligned}$$

where  $\sum^{\dagger}$  denotes that the sum is restricted to vectors failing coprimality at every prime  $p$ ; i.e. the vector  $(v_p(\bar{d}_j \bar{q}_j \bar{e}_j \bar{e}'_j))_{j=1}^{r'}$  has either no or at least two non-zero entries.

Our next aim is to bound this second factor from above. We will do this by writing it as an Euler product and analysing contributions for each prime factor separately. The saving in the bound will come from the factor  $\alpha_{\mathbf{h}'}$  in  $\eta$ . The remaining factors may be bounded trivially by

$$\left| \mu(p^{e_j}) \mu(p^{e'_j}) p^{-e_j z_{j,1}} p^{-e'_j z_{j,2}} g_{i_j}(p^{d_j}) p^{-d_j z_{j,3}} \tau(p^{q_j})^{n_{i_j}} p^{-q_j z_{j,4}} \right| \leq 2^{d_j n_{i_j}} (q_j + 1)^{n_{i_j}}. \quad (9.9)$$

In order to turn the sum over  $\bar{\mathbf{d}}, \bar{\mathbf{q}}, \bar{\mathbf{e}}, \bar{\mathbf{e}}'$  into one that directly runs over the entries of  $\alpha_{\mathbf{h}'}$ , note that any integer  $k_j$  may be factorised as  $\hat{\Delta}_j$  in at most  $\tau_5(k_j)$  ways, corresponding to the five factors  $d_j$ ,  $q_j$ ,  $\gcd(e_j, e'_j)$ ,  $e_j / \gcd(e_j, e'_j)$  and  $e'_j / \gcd(e_j, e'_j)$ . We will employ the crude bound  $\tau_5(p^{a_j}) \ll a_j^4$ . Let  $n(\mathbf{x})$  denote the number of non-zero components of  $\mathbf{x} \in \mathbb{Z}^{r'}$ . Then the previous inequality implies

$$\sum_{\substack{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}' \\ n(\mathbf{d} + \mathbf{q} + \max(\mathbf{e}, \mathbf{e}')) \geq 2}} \eta(p^{\mathbf{d}}, p^{\mathbf{q}}, p^{\mathbf{e}}, p^{\mathbf{e}'}) \ll \sum_{\substack{\mathbf{a} \in \mathbb{Z}_{\geq 0}^{r'} \\ n(\mathbf{a}) \geq 2}} \alpha_{\mathbf{h}'}(p^{a_1}, \dots, p^{a_{r'}}) \prod_{j=1}^{r'} C^{a_j} (a_j + 1)^C,$$

for some absolute positive constant  $C$ . Assuming  $p > w(T)$  for sufficiently large  $T$  and introducing the variable  $J = \max_{j \neq j'} (a_j + a_{j'})$ , the third case of (5.6) shows that this in turn is bounded by

$$\ll \sum_{J \geq 2} p^{-J} C^{r'J} J^{r'C} \ll \frac{1}{p^2}.$$

Recall that all components of  $\mathbf{d}, \mathbf{q}, \mathbf{e}$  and  $\mathbf{e}'$  are composed only of prime factors larger than  $w(T)$ . In total, we deduce that

$$\begin{aligned} \sum_{\substack{(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \\ (\bar{\mathbf{d}}, \bar{\mathbf{q}}, \bar{\mathbf{e}}, \bar{\mathbf{e}}') \neq (1, 1, 1, 1)}} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') &\leq \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \left( \prod_{p > w(T)} \left( 1 + O\left(\frac{1}{p^2}\right) \right) - 1 \right) \\ &\leq \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}') \left( \sum_{m > w(T)} m^{-3/2} \right) \\ &\ll w(T)^{-1/2} \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'). \end{aligned}$$

Thus, for the treatment of the left hand side of the expression from the lemma, it remains to prove (9.8). Employing (9.9) and the bound on  $\tau_5$  another time, we turn the sum into one that only involves  $\alpha_{\mathbf{h}'}$  and may be estimated using the second part of (5.6). Introducing the variable  $Q = \max_j(a_j)$ , we have

$$\begin{aligned} \sum'_{p^{\mathbf{d}}, p^{\mathbf{q}}, p^{\mathbf{e}}, p^{\mathbf{e}'}} \eta(p^{\mathbf{d}}, p^{\mathbf{q}}, p^{\mathbf{e}}, p^{\mathbf{e}'}) &= 1 + O \left( \sum_{\substack{\mathbf{a} \in \mathbb{Z}_{\geq 0}^{r'} \\ n(\mathbf{a})=1}} \alpha_{\mathbf{h}'}(p^{a_1}, \dots, p^{a_{r'}}) \prod_{j=1}^{r'} C^{a_j} (a_j + 1)^C \right) \\ &= 1 + O \left( \sum_{Q \geq 1} p^{-Q} C^Q (Q + 1)^{C+4} \right) \\ &= 1 + O(p^{-1}). \end{aligned}$$

This completes the proof of the estimate

$$\begin{aligned} \sum_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'}) J_1 \dots J_{r'} &= (1 + o(1)) \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \alpha_{\mathbf{h}'}(\tilde{\Delta}_1, \dots, \tilde{\Delta}_{r'}) J_1 \dots J_{r'} \\ &= (1 + o(1)) \sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \prod_{j=1}^{r'} \frac{J_j}{\tilde{\Delta}_j}. \end{aligned}$$

To complete the proof of the lemma, we need to show that in fact

$$\sum'_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \prod_{j=1}^{r'} \frac{J_j}{\tilde{\Delta}_j} = (1 + o(1)) \sum_{\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}'} \prod_{j=1}^{r'} \frac{J_j}{\tilde{\Delta}_j}.$$

This follows by arguing as above when  $\eta(\mathbf{d}, \mathbf{q}, \mathbf{e}, \mathbf{e}')$  is redefined to equal  $\prod_{j=1}^{r'} J_j / \tilde{\Delta}_j$ , and when taking into account that the product  $(p^{a_1} \dots p^{a_{r'}})^{-1}$ , which replaces  $\alpha_{\mathbf{h}'}(p^{a_1}, \dots, p^{a_{r'}})$ , trivially satisfies the bounds (5.6).  $\square$

*Proof of Proposition 9.2.* A slight adaptation of [18, Lemma 9.9] yields the following. Let  $\Delta : \mathbb{Z} \rightarrow \mathbb{Z}$  denote the polynomial

$$\Delta(m) = \prod_{1 \leq j < j' \leq d} (Wm + A_{i_j} - A_{i_{j'}}).$$

Suppose  $\sigma : \{-T', \dots, T'\} \rightarrow \mathbb{R}$  satisfies the two conditions  $\sigma(0) = O(T'^{1/q})$  and

$$\sigma(m) = \exp \left( \sum_{p > w(T), p | \Delta(m)} O_T(p^{-1/2}) \right)$$

for  $m \neq 0$ . Then  $\mathbb{E}_{0 \leq m \leq \tilde{T}} \sigma^q(m) \ll_q 1$ .

Whenever the collection of  $a_j$  contains two identical elements, then  $\sigma(0)$  appears in the bound we seek to establish. Following [17, 18] closely, we use the fact that  $\sigma(0)$  may be chosen to be rather large in order to handle this case. More precisely, it follows from Hölder's inequality and the fact that  $r_{i,\text{res}}$  satisfies part (b) of Definition 7.1, that

$$\frac{1}{T'} \sum_{m \in I} \prod_{j=1}^d \varphi_{i_j}(T; A_{i_j}) \nu_{i_j}^{(T)}(W(m + a_j) + A_{i_j}) \nu_{i_j, \text{sieve}}^{(T)}(W(m + a_j) + A_{i_j}) \ll_{c,d} T^{cd}.$$

See [17, §9] and [26, §7] for details. Choosing  $c = 1/(2qd)$  ensures that the value on the right hand side is of order  $o((T/W)^{1/q})$ , so that we may set  $\sigma(0) = T^{1/2q}$ .

In the remaining case where the  $a_j$  are pairwise distinct, the system of linear forms is less degenerate and we may employ the same techniques used to prove Proposition 9.1. The key observation is that whenever a prime  $p$  divides two distinct polynomials  $W(m + a_j) + A_{i_j}$  and  $W(m + a_{j'}) + A_{i_{j'}}$  at  $m$ , then it divides  $W(a_j - a_{j'}) + A_{i_j} - A_{i_{j'}}$ . This provides sufficient information to handle the divisor densities  $\alpha(p^{a_1}, \dots, p^{a_d})$  that occur. See [26, §7] for details.  $\square$

## 10. CONCLUSION OF THE PROOF

We have now everything in place in order to complete the proof of Theorem 5.2. Recall from (5.1) that

$$N(T) = \sum_{\substack{\mathbf{u} \in \mathbb{Z}^s \cap T\mathfrak{K} \\ \mathbf{u} \equiv \mathbf{a} \pmod{M}}} \prod_{i=1}^r R_i(f_i(\mathbf{u})),$$

where  $R_i(m) = R_i(m; \mathfrak{X}_i, \mathbf{b}_i; M)$  is given by Definition 5.1 for non-zero  $m \in \mathbb{Z}$ . Here,  $M \in \mathbb{Z}_{>0}$ ,  $\mathbf{a} \in (\mathbb{Z}/M\mathbb{Z})^s$  and  $\mathbf{b}_i \in (\mathbb{Z}/M\mathbb{Z})^{n_i}$  for  $1 \leq i \leq r$ . Moreover,  $\mathfrak{X}_i \subset \mathfrak{D}_{i,+}$  is a cone for which the bounded set  $\mathfrak{X}_i \cap \mathfrak{D}_{i,+}^\epsilon(1)$  has an  $(n_i - 1)$ -Lipschitz parametrisable boundary, unless it is empty. Finally,  $\mathfrak{K} \subset \mathbb{R}^s$  is a convex bounded set.

Recall the definition (6.2) of  $\mathcal{A}_i$ , for each  $1 \leq i \leq r$ . Let

$$\mathscr{W} = \left\{ \mathbf{u}_0 \in (\mathbb{Z}/W\mathbb{Z})^s : \begin{array}{l} f_i(\mathbf{u}_0) \in \mathcal{A}_i \text{ for } i = 1, \dots, r \\ \mathbf{u}_0 \equiv \mathbf{a} \pmod{M} \end{array} \right\},$$

where  $W$  is given by (6.1) and is divisible by  $M$ . For each  $p < w(T)$  we define the corresponding set

$$\mathcal{W}_p = \left\{ \mathbf{u}_0 \in (\mathbb{Z}/p^{v_p(W)}\mathbb{Z})^s : \begin{array}{l} v_p(f_i(\mathbf{u}_0)) < v_p(W)/3 \text{ for } i = 1, \dots, r \\ \mathbf{u}_0 \equiv \mathbf{a} \pmod{p^{v_p(M)}} \\ \varrho_i(p^{v_p(W)}, f_i(\mathbf{u}_0); p^{v_p(M)}) > 0 \end{array} \right\},$$

where we recall that  $\varrho_i(p^{v_p(W)}, f_i(\mathbf{u}_0); p^{v_p(M)})$  also depends on  $\mathbf{b}_i$ . Define the functions  $R'_i$  as in (8.1) and recall the exceptional set  $\mathcal{S}_{C_1, T}$  from Definition 7.5. We note that

$$R'_i(m) = \sum_{A \in \mathcal{A}_i} \mathbf{1}_{m \equiv A \pmod{W}} R'_i(m),$$

for  $1 \leq i \leq r$ . Indeed, suppose  $|m| \notin \mathcal{S}_{C_1, T}$  with  $R'_i(m) \neq 0$ . Then it is clear that the reduction of  $m$  modulo  $W$  must belong to  $\mathcal{A}_i$ . By Proposition 8.2, it therefore suffices to obtain an asymptotic for

$$N'(T) = \sum_{\mathbf{u}_0 \in \mathcal{W}} \sum_{\substack{\mathbf{u}_1 \in \mathbb{Z}^s \\ W\mathbf{u}_1 + \mathbf{u}_0 \in T\mathfrak{K}}} \prod_{i=1}^r R'_i(f_i(W\mathbf{u}_1 + \mathbf{u}_0)). \quad (10.1)$$

Next, let  $i \in \{1, \dots, r\}$ . We have  $f_i(W\mathbf{u}_1 + \mathbf{u}_0) = Wf_i(\mathbf{u}_1) + f_i(\mathbf{u}_0)$ , since each  $f_i$  is a linear form. Let  $0 < A'_i(\mathbf{u}_0) < W$  be such that  $A'_i(\mathbf{u}_0) \equiv f_i(\mathbf{u}_0) \pmod{W}$ . We proceed to define a linear polynomial  $h_i \in \mathbb{Z}[\mathbf{u}]$  via

$$Wf_i(\mathbf{u}) + f_i(\mathbf{u}_0) = Wh_i(\mathbf{u}) + A'_i(\mathbf{u}_0).$$

Note that  $h_i$  may be inhomogeneous.

For any fixed residue  $\mathbf{u}_0 \in \mathcal{W}$ , let  $\mathfrak{K}_{\mathbf{u}_0, T}$  be the set of  $\mathbf{u}_1 \in \mathbb{R}^s$  for which  $W\mathbf{u}_1 + \mathbf{u}_0 \in T\mathfrak{K}$ . Thus,  $W\mathfrak{K}_{\mathbf{u}_0, T} + \mathbf{u}_0 = T\mathfrak{K}$ . We proceed to split  $\mathfrak{K}_{\mathbf{u}_0, T}$  into regions on which the sign of  $Wh_i(\mathbf{u}_1) + A'_i$  is constant for  $1 \leq i \leq r$ . Thus, for  $\epsilon = (\epsilon_1, \dots, \epsilon_r) \in \{\pm 1\}^r$  let

$$\mathfrak{K}_{\mathbf{u}_0, T}(\epsilon) = \{\mathbf{u}_1 \in \mathfrak{K}_{\mathbf{u}_0, T} : \mathbf{f}(W\mathbf{u}_1 + \mathbf{u}_0) \in \mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r}\},$$

where  $\mathbb{R}_\epsilon = \{x \in \mathbb{R} : \epsilon x > 0\}$ . Note that this is a finite union of convex subsets of  $[-CT/W, CT/W]^s$  for some absolute constant  $C > 0$ . Furthermore

$$\text{vol}(\mathfrak{K}_{\mathbf{u}_0, 1}(\epsilon)) = \frac{\text{vol}(\mathfrak{K} \cap \mathbf{f}^{-1}(\mathbb{R}_{\epsilon_1} \times \dots \times \mathbb{R}_{\epsilon_r}))}{W^s}.$$

Since  $|\mathbf{f}(T\mathfrak{K})| \leq T$ , it follows that  $|Wh_i(\mathfrak{K}_{\mathbf{u}_0, T}) + A'_i(\mathbf{u}_0)| \leq T$  for each  $\mathbf{u}_0 \in \mathcal{W}$ .

The existence of a simultaneous pseudorandom majorant for each collection of functions

$$\tilde{R}_i : m \mapsto \left( \frac{\varrho_i(W, f_i(\mathbf{u}_0); M)}{W^{n_i-1}} \right)^{-1} R'_i(W\epsilon_i m + A'_i(\mathbf{u}_0))$$

defined on the range  $\{m : 0 < Wm + \epsilon_i A'_i(\mathbf{u}_0) \leq T\}$  was established in Sections 8 and 9. This existence allows us to employ the generalised von Neumann theorem [18, Prop. 7.1] to deduce that the sum over  $\mathbf{u}_1$  in (10.1) is equal to

$$T^s \sum_{\epsilon \in \{\pm 1\}^r} \text{vol}(\mathfrak{K}_{\mathbf{u}_0, 1}(\epsilon)) \prod_{i=1}^r \frac{\varrho_i(W, f_i(\mathbf{u}_0); M)}{W^{n_i-1}} \kappa_i^{\epsilon_i}(\mathfrak{X}_i) + o\left(\frac{T^s}{W^s}\right),$$



provided that for each  $\mathbf{u}_0 \in \mathscr{W}$  the normalised representation function  $\tilde{R}_i(m)$  satisfies

$$\max_{1 \leq i \leq r} \|\tilde{R}_i - \kappa_i^{\epsilon_i}(\mathfrak{X}_i)\|_{U^{r-1}} = o(1).$$

The latter, however, follows from the inverse theorem [21] for the Gowers uniformity norms from Proposition 6.3 and the bound

$$\mathbb{E}_{|m| < T/W} R_i(Wm + A) \mathbf{1}_{Wm + A \in \mathcal{S}_{C_1, T}} \ll (\log T)^{-C_1/4},$$

provided by (8.2).

Let

$$\mathfrak{S}(T) = \frac{1}{W^s} \sum_{\mathbf{u}_0 \in \mathscr{W}} \prod_{i=1}^r \prod_{p \leq w(T)} \frac{\varrho_i(p^{\alpha(p)}, f_i(\mathbf{u}_0); p^{v_p(M)})}{p^{\alpha(p)(n_i-1)}},$$

with  $\alpha(p) = v_p(W)$ . We conclude that

$$N'(T) = \beta_\infty \mathfrak{S}(T) T^s + o(T^s),$$

with  $\beta_\infty$  as in the statement of Theorem 5.2. It therefore remains to analyse  $\mathfrak{S}(T)$ . An application of the Chinese remainder theorem yields

$$\mathfrak{S}(T) = \prod_{p \leq w(T)} \frac{1}{p^{s\alpha(p)}} \sum_{\mathbf{u}_0 \in \mathscr{W}_p} \prod_{i=1}^r \frac{\varrho_i(p^{\alpha(p)}, f_i(\mathbf{u}_0); p^{v_p(M)})}{p^{\alpha(p)(n_i-1)}}.$$

Let us define  $\varepsilon_p$  via

$$\beta_p = \frac{1}{p^{s\alpha(p)}} \sum_{\mathbf{u}_0 \in \mathscr{W}_p} \prod_{i=1}^r \frac{\varrho_i(p^{\alpha(p)}, f_i(\mathbf{u}_0); p^{v_p(M)})}{p^{\alpha(p)(n_i-1)}} + \varepsilon_p,$$

where  $\beta_p$  is as in the statement of Theorem 5.2. In order to complete the proof, it remains to check that  $\varepsilon_p$  is sufficiently small to be able to conclude that

$$\prod_{p < w(T)} (\beta_p - \varepsilon_p) = \prod_{p < w(T)} \beta_p + o(1),$$

as  $T \rightarrow \infty$ . This will certainly suffice, since Proposition 5.5 implies that

$$\prod_{p > w(T)} \beta_p = 1 + o(1),$$

as  $T \rightarrow \infty$ . Recalling that  $w(T) = \log \log T$ , it will be enough to show that  $\varepsilon_p \ll (\log T)^{-C}$  for some absolute constant  $C > 0$ .

For this we shall apply Lemma 3.4 to  $\gamma(p^m, A; p^\ell) = \varrho_i(p^m, f_i(\mathbf{u}_0); p^{v_p(M)})$  for  $m \geq \alpha(p)$ . This yields

$$\begin{aligned} \varepsilon_p &= \lim_{m \rightarrow \infty} \frac{1}{p^{ms}} \sum_{\substack{\mathbf{u} \in \mathcal{U}_m^* \\ \mathbf{u} \equiv \mathbf{a} \pmod{M}}} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}); p^{v_p(M)})}{p^{m(n_i-1)}} \\ &\leq \lim_{m \rightarrow \infty} \frac{1}{p^{ms}} \sum_{\mathbf{u} \in \mathcal{U}_m^*} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}))}{p^{m(n_i-1)}}, \end{aligned}$$

where

$$\mathcal{U}_m^* = \left\{ \mathbf{u} \in (\mathbb{Z}/p^m\mathbb{Z})^s : \max_{1 \leq i \leq r} v_p(f_i(\mathbf{u})) \geq \min\{\alpha(p)/3, m\} \right\}.$$

Proceeding as in the analysis of  $\beta_p$  for large  $p$  in the proof of Proposition 5.5, we obtain for  $m \geq \alpha(p)$  the bound

$$\begin{aligned} \frac{1}{p^{ms}} \sum_{\mathbf{u} \in \mathcal{U}_m^*} \prod_{i=1}^r \frac{\varrho_i(p^m, f_i(\mathbf{u}))}{p^{m(n_i-1)}} &\ll \frac{1}{p^{ms}} \sum_{\mathbf{u} \in \mathcal{U}_m^*} \prod_{i=1}^r \min\{v_p(f_i(\mathbf{u})) + 1, m\}^{n_i} \\ &\leq \sum_{\substack{\mathbf{k} \in \mathbb{Z}_{\geq 0}^r \\ \max_i k_i \geq \alpha(p)/3}} \alpha_{\mathbf{f}}(p^{k_1}, \dots, p^{k_r}) \prod_{i=1}^r k_i^{n_i} \\ &\ll p^{-\alpha(p)/3} \\ &\leq (\log T)^{-C_1/4}. \end{aligned}$$

This completes the proof of Theorem 5.2.

## REFERENCES

- [1] T.D. Browning and D.R. Heath-Brown, Quadratic polynomials represented by norm forms. *GAF* **22** (2012), 1124–1190.
- [2] T.D. Browning, L. Matthiesen and A.N. Skorobogatov, Rational points on pencils of conics and quadrics with many degenerate fibres. *Annals of Math.* **180** (2014), 381–402.
- [3] J.-L. Colliot-Thélène, Surfaces rationnelles fibrées en coniques de degré 4. *Séminaire de théorie des nombres, Paris 1988–1989*, 43–55, Progr. Math. **91**, Birkhäuser, 1990.
- [4] J.-L. Colliot-Thélène. Points rationnels sur les fibrations. *Higher dimensional varieties and rational points (Budapest, 2001)*, 171–221, Springer-Verlag, 2003.
- [5] J.-L. Colliot-Thélène and P. Salberger, Arithmetic on some singular cubic hypersurfaces. *Proc. London Math. Soc.* **58** (1989), 519–549.
- [6] J.-L. Colliot-Thélène and J.J. Sansuc, La  $R$ -équivalence sur les tores. *Ann. Sci. École Norm. Sup.* **10** (1977), 175–229.
- [7] J.-L. Colliot-Thélène and J.-J. Sansuc, La descente sur les variétés rationnelles, II. *Duke Math. J.* **54** (1987), 375–492.
- [8] J.-L. Colliot-Thélène and P. Swinnerton-Dyer, Hasse principle and weak approximation for pencils of Severi–Brauer and similar varieties. *J. reine angew. Math.* **453** (1994), 49–112.

- [9] J.-L. Colliot-Thélène, D. Harari and A.N. Skorobogatov, Valeurs d'un polynôme à une variable représentés par une norme. *Number theory and algebraic geometry*, 69–89, London Math. Soc. Lecture Note Ser. **303** Camb. Univ. Press, 2003.
- [10] J.-L. Colliot-Thélène, J.-J. Sansuc and P. Swinnerton-Dyer, Intersections of two quadrics and Châtelet surfaces, I. *J. reine angew. Math.* **373** (1987), 37–107.
- [11] J.-L. Colliot-Thélène, J.-J. Sansuc and P. Swinnerton-Dyer, Intersections of two quadrics and Châtelet surfaces, II. *J. reine angew. Math.* **374** (1987), 72–168.
- [12] J.-L. Colliot-Thélène, A.N. Skorobogatov and P. Swinnerton-Dyer, Rational points and zero-cycles on fibred varieties: Schinzel's hypothesis and Salberger's device. *J. reine angew. Math.* **495** (1998), 1–28.
- [13] U. Derenthal, A. Smeets and D. Wei, Universal torsors and values of quadratic polynomials represented by norms. *Math. Annalen* **361** (2015), 1021–1042.
- [14] P. Erdős, On the sum  $\sum_{k=1}^x d(f(k))$ . *J. London Math. Soc.* **27** (1952), 7–15.
- [15] D.A. Goldston and C.Y. Yıldırım, Higher correlations of divisor sums related to primes. III. Small gaps between primes. *Proc. Lond. Math. Soc.* **95** (2007), 653–686.
- [16] D.A. Goldston, J. Pintz and C.Y. Yıldırım, Primes in tuples. I. *Annals of Math.* **170** (2009), 819–862.
- [17] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions. *Annals of Math.* **167** (2008), 481–547.
- [18] B. Green and T. Tao, Linear equations in primes. *Annals of Math.* **171** (2010), 1753–1850.
- [19] B. Green and T. Tao, The quantitative behaviour of polynomial orbits on nilmanifolds. *Annals of Math.* **175** (2012), 465–540.
- [20] B. Green and T. Tao, The Möbius function is strongly orthogonal to nilsequences. *Annals of Math.* **175** (2012), 541–566.
- [21] B. Green, T. Tao and T. Ziegler, An inverse theorem for the Gowers  $U_{s+1}[N]$ -norm. *Annals of Math.* **176** (2012), 1231–1372.
- [22] Y. Harpaz, A.N. Skorobogatov and O. Wittenberg, The Hardy–Littlewood conjecture and rational points. *Compositio Math.* **150** (2014), 2095–2111.
- [23] D.R. Heath-Brown and A.N. Skorobogatov, Rational solutions of certain equations involving norms. *Acta Math.* **189** (2002), 161–177.
- [24] E. Landau, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale*. Teubner Verlag, Leipzig, 1918.
- [25] D.A. Marcus, *Number fields*. Springer-Verlag, 1977.
- [26] L. Matthiesen, Correlations of the divisor function. *Proc. London Math. Soc.* **104** (2012), 827–858.
- [27] L. Matthiesen, Linear correlations amongst numbers represented by positive definite binary quadratic forms. *Acta Arith.* **154** (2012), 235–306.
- [28] L. Matthiesen, Correlations of representation functions of binary quadratic forms. *Acta Arith.* **158** (2013), 245–252.
- [29] H.L. Montgomery and R.C. Vaughan, *Multiplicative Number Theory, I. Classical Theory*. Camb. Univ. Press, 2007.
- [30] J. Neukirch, *Algebraic number theory*. Springer-Verlag, 1991.
- [31] D. Schindler and A.N. Skorobogatov, Norms as products of linear polynomials. *J. London Math. Soc.* **89** (2014), 559–580.
- [32] J.-P. Serre, *A course in arithmetic*. Springer-Verlag, 1996.
- [33] P. Shiu, A Brun–Titchmarsh theorem for multiplicative functions. *J. reine angew. Math.* **313** (1980), 161–170.
- [34] M. Swarbrick Jones, A note on a theorem of Heath-Brown and Skorobogatov. *Q. J. Math.* **64** (2013), 1239–1251.
- [35] D. Wei. On the equation  $N_{K/k}(\Xi) = P(t)$ . *Proc. London Math. Soc.* **109** (2014), 1402–1434.

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL, BS8 1TW  
*E-mail address:* `t.d.browning@bristol.ac.uk`

KTH, DEPARTMENT OF MATHEMATICS, 10044 STOCKHOLM  
*E-mail address:* `lilian.matthiesen@math.kth.se`